

# Essays on Plato's Psychology

Edited by  
Ellen Wagner



LEXINGTON BOOKS  
*Landham • Boulder • New York • Oxford*

MAIN  
B  
698  
S7  
E7  
001

LEXINGTON BOOKS

Published in the United States of America  
by Lexington Books  
4720 Boston Way, Lanham, Maryland 20706

12 Hid's Copse Road  
Cumnor Hill, Oxford OX2 9JJ, England

Copyright © 2001 by Ellen Wagner

*All rights reserved.* No part of this publication may be reproduced,  
stored in a retrieval system, or transmitted in any form or by any  
means, electronic, mechanical, photocopying, recording, or otherwise,  
without the prior permission of the publisher.

British Library Cataloguing in Publication Information Available

**Library of Congress Cataloging-in-Publication Data**

Essays on Plato's psychology / edited by Ellen Wagner.

p. cm.

Includes bibliographical references and index.

ISBN 0-7391-0256-7 (alk. paper) — ISBN 0-7391-0258-3 (pbk. : alk. paper)

1. Plato. 2. Soul. I. Wagner, Ellen, 1949-

B398.S7 E77 2001  
128'.1'092—dc21

2001029540

Printed in the United States of America

 The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI/NISO Z39.48—1992.

*To the memory of  
Pauline Anna Cook Wendler  
1903-1991  
and for  
Paula Anna Wendler  
with love*

CC 4.12.02

# Contents

Acknowledgments	ix
Introduction	1
<b>I. The Nature of the Soul</b>	
1 Fallacies in the <i>Phaedo</i> Again <i>Cass Weller</i>	35
2 The Arguments in the <i>Phaedo</i> Concerning the Thesis That the Soul Is a <i>Harmonia</i> <i>C. C. W. Taylor</i>	51
3 Supervenience and the Thesis That the Soul Is a <i>Harmonia</i> <i>Ellen Wagner</i>	69
<b>II. The Tripartite Soul</b>	
4 Plato's Theory of Human Motivation <i>John M. Cooper</i>	91
5 Plato's Analogy of Soul and State <i>Nicholas D. Smith</i>	115
6 Simple Souls <i>Christopher Shields</i>	137
7 The Analogy of City and Soul in Plato's <i>Republic</i> <i>Bernard Williams</i>	157

8 Inside and Outside the <i>Republic</i> <i>Jonathan Lear</i>	169
9 Akrasia and Agency in Plato's <i>Laws</i> and <i>Republic</i> <i>Christopher Bobonich</i>	203
<b>III. Arguments for Immortality of the Soul</b>	
10 The Soul and Immortality in Plato's <i>Phaedo</i> <i>David Bostock</i>	241
11 Plato's "Cyclical" Argument Recycled <i>David Gallop</i>	263
12 The Final Proof of the Immortality of the Soul in Plato's <i>Phaedo</i> 102a-107a <i>Dorothea Frede</i>	281
13 A Defense of Plato's Argument for the Immortality of the Soul at <i>Republic</i> X 608c-611a <i>Eric A. Brown</i>	297
14 The Argument for Immortality in Plato's <i>Phaedrus</i> <i>Thomas M. Robinson</i>	323
15 Immortality and the Nature of the Soul in the <i>Phaedrus</i> <i>Richard Bett</i>	335
Selected Bibliography	363
Index	371
Index Locorum	375
About the Contributors	383

## Acknowledgments

A number of friends and colleagues have helped me in the preparation of this book. I would like to thank especially Christopher Shields and Nicholas D. Smith for valuable advice at several points, and Hugh Benson for suggestions on the project itself. Rosa De Jorio inspired me to propose the book. I am grateful to the staff at Lexington Books, especially Serena Leigh, Ginger Strader, and Melissa McNitt; to the University of North Florida for awarding me a research grant; and to Marianne Roberts for her invaluable computer expertise. Special appreciation is due Ashley Godeaux for his love and support throughout the editorial process.

In addition, I thank the following people for their assistance and their permission to use the materials printed or reprinted in this book:

Cass Weller: "Fallacies in the *Phaedo* Again" first appeared in *Archiv für Geschichte der Philosophie* 77 (1995): 121-34. Reprinted with permission of Walter de Gruyter GmbH & Co. KG and the author.

C. C. W. Taylor: "The Arguments in the *Phaedo* Concerning the Thesis That the Soul Is a *Harmonia*" first appeared in *Essays in Ancient Greek Philosophy*, Vol. II, ed. John P. Anton and Anthony Preus (Albany, N.Y.: State University of New York Press, 1983), 217-32.

Ellen Wagner: "Supervenience and the Thesis That the Soul Is a *Harmonia*" first appeared in the *Southwest Philosophy Review* 16 (2000): 1-20. Reprinted with permission of the author and thanks to the *Southwest Philosophy Review*.

John Cooper: "Plato's Theory of Human Motivation" first appeared in *History of Philosophy Quarterly* 1 (1984): 3-24. Reprinted with permission of *History of Philosophy Quarterly* and the author.

Nicholas D. Smith: "Plato's Analogy of Soul and State" first appeared in *Journal of Ethics* 3 (1999): 31-49. Reprinted with kind permission of Kluwer Academic Publishers and the author.

Christopher Shields: "Simple Souls" is printed with permission of the author.

Bernard Williams: "The Analogy of City and Soul in Plato's *Republic*" first appeared in *Exegesis and Argument*, ed. E. N. Lee, A. P. D. Mourelatos, and R. M. Rorty (Assen: Van Gorcum & Company B. V., 1973), 196-206. Reprinted with permission of the author.

Jonathan Lear: "Inside and Outside the *Republic*" first appeared in *Phronesis* 37 (1992): 184-215. Reprinted with permission of Brill Academic Publishers and the author.

Christopher Bobonich: "Akrasia and Agency in Plato's *Laws* and *Republic*" first appeared in *Archiv für Geschichte der Philosophie* 76 (1994): 3-36. Reprinted with permission of Walter de Gruyter GmbH & Co. KG and the author.

David Bostock: "The Soul and Immortality in Plato's *Phaedo*." Originally published in *Plato's Phaedo* by David Bostock (1986), © David Bostock 1986. Published in a revised form, followed here, in *Plato 2: Ethics, Politics, Religion, and the Soul*, edited by Gail Fine (1999). Reprinted with permission of Oxford University Press and the author.

David Gallop: "Plato's 'Cyclical' Argument Recycled" first appeared in *Phronesis* 27 (1982): 207-22. Reprinted with permission of Brill Academic Publishers and the author.

Dorothea Frede: "The Final Proof of the Immortality of the Soul in Plato's *Phaedo* 102a-107a" first appeared in *Phronesis* 23 (1978): 27-41. Reprinted with permission of Brill Academic Publishers and the author.

Eric A. Brown: "A Defense of Plato's Argument for the Immortality of the Soul at *Republic* X 608c-611a" is reprinted from *APEIRON: a journal for ancient philosophy and science* 30 (1997) by permission of Academic Printing and Publishing and the author.

Richard Bett: "The Argument for Immortality in Plato's *Phaedrus*" first appeared in *Phronesis* 31 (1986): 1-26. Reprinted with permission of Brill Academic Publishers and the author.

T. M. Robinson: "The Argument for Immortality in Plato's *Phaedrus*" first appeared in *Essays in Ancient Greek Philosophy*, ed. John P. Anton and George L. Kustas (Albany, N.Y.: State University of New York Press, 1971), 345-53. Reprinted with permission of the State University of New York Press and the author.

## Introduction

This anthology collects in a single volume some of the finest scholarly work on three central topics of Plato's psychology or theory of the soul. The articles presented here are seminal explorations of (1) the nature of soul in Plato, (2) the famous question of his tripartite division of the soul (notably in the *Republic*), and (3) the strand of argumentation for the immortality of soul that is woven through several dialogues. Authors whose work is represented have collectively produced now-classic translations of, and commentaries on, the dialogues, as well as both analyses of specific arguments and discussions of the works' overarching strategies and dramatic structures.

Given the recent explosion of Plato scholarship, it is not surprising that the psychology should have received such a wealth of attention; what is surprising is that there has not ever previously been an anthology devoted to this subject.<sup>1</sup> This state of affairs may be due to two factors: the overwhelming attention focused on such extremely difficult topics as Plato's theory of Forms; and more importantly, the development of new interpretive approaches and increased interest in rethinking the dialogues. At any rate, this collection is a reflection of the remarkable vigor and excellence of Plato studies today both in general and on his psychology, and should spur additional explorations of the latter.

When philosophers think of psychological theories originating with the ancient Greeks, it is Aristotle's *De Anima* that immediately comes to mind, with its methodical investigations of the soul's faculties and activities. Yet Plato's psychology, which is comparably rich and insightful, also offers us a compelling account of human motivation, internal conflict, and character. In the words of a recent commentator, "[f]ar from

being an ancient curiosity . . . Plato's theory of the psyche, largely coherent, supported by subtle argument, and possessed of considerable folk-psychological plausibility, is among the greatest philosophies of mind, and one from which we can still learn.”<sup>2</sup> His portrayals of the development of both fine and corrupt characters in *Republic* 7, 8, and 9, for example, reflect a shrewd, accurate understanding of human character and motivation. Arguments for the modularity of the soul and explanations of how conflicts arise within it, as well as an educational program for developing the correct order in the soul, remind us of just how difficult it is to master one's own appetites and emotions, while they acknowledge the need for all of the soul's capacities to be developed effectively. One recognizes in Plato a careful observer of human nature who designs a complex educational program to elicit and to shape the best in that nature.

Many of the questions about the soul that Plato tried to answer in the dialogues are also those that contemporary philosophers of mind have renewed in modern terminology: How is the mind different in substance, if at all, from the body? What is its precise relation to the body? Is the soul (mind) simple or complex, and in what sense is it so?<sup>3</sup> Modern philosophers of mind find much of interest in Plato's attempts to refute alternative conceptions of the soul, notably the *harmonia* thesis in the *Phaedo*, as well as in his descriptions of the components of the soul in the *Republic* and *Phaedrus*, in which the soul is composed of parts, frequently characterized as functional parts, that anticipate modular theories of mind.<sup>4</sup> We might also note the recent surge of interest in the hard problem of consciousness—that is, explaining how neural events yield corresponding conscious experiences—which is at least partially responsible for the legitimization of consciousness studies within philosophy of mind.<sup>5</sup> Just as Plato wondered in what ways the soul and body were in fact different entities, and if so how the soul or mind achieved its connection to the body, so do contemporary philosophers of mind try to formulate views on supervenience of the mind on the body, mind as a set of functions of the body, and the production of qualia. In addition, psychologists today bear witness to Plato's inspiration: it is clear that Freud owed a debt to Plato's tripartite division of the soul in his own formulation of id, ego, and superego. Frequently issues of psychological structure originate historically to a significant extent in Plato's questions and hypotheses.

Of course, it is also a matter of interest to many people whether the soul survives the death of the body. Hence, Plato's inventive arguments for immortality, mostly based on observations about cognition and other

empirical facts about the nature of living things, are relevant. In a material world, the attempt to prove the existence of an immaterial soul that survives our physical death is a difficult endeavor; Plato's arguments demonstrate this clearly.<sup>6</sup> Yet although the term “Platonic dualism” is commonly used to refer to a strong dualism of material body and independent, separate immaterial soul, this general rubric belies the subtlety of his psychology. Arguing for the pre- and postmortem existence of the soul, he is nonetheless aware of its complex interaction with the body. His formulations of the relevant issues should be of interest to anyone who seriously considers the possibility of a soul's surviving the body's death.

The importance of Plato's psychology in his philosophy extends beyond its immediate borders, since this theory is a central component of his metaphysics, epistemology, and ethics. The soul bridges the gap between the sensible world and the realm of the Forms, making knowledge possible as well as explaining the nature of moral character, its perfection, and its degeneration. Indeed, the twin pillars of Plato's philosophy were famously said by Cornford to be his beliefs in the world of the Forms and in the immortality of the soul.<sup>7</sup> In order to understand Plato's philosophical thought as a whole, then, one must have a firm grasp of his views of the soul. What we have inherited from Plato is more than a quaint ancient theory: it is an insightful set of theoretical and empirical observations well worth our careful attention.

Before we begin a systematic study of the psychology, however, we must consider the rich and complex background of the ideas Plato articulated about the soul.<sup>8</sup> Studying the psychology, one notices immediately its seeming internal contradictions. In different dialogues Plato's Socrates argues for (1a) a perfectly simple soul; (1b) a soul with three distinct components or parts; (2a) immortality for the reasoning part only; (2b) the immortality of all three components—and immortality on different grounds in each of numerous arguments; (3a) the soul's natural ability to rule the body; and (3b) its helpless imprisonment in that body. One way to make sense of the disparate elements that Plato introduces in his theory of the soul is to acknowledge his inheritance from traditional Homeric views, as well as the Orphic-Pythagorean tenets which he adopted. I turn, therefore, to considering both the Homeric conception of the soul and the Orphic-Pythagorean beliefs that play important roles in his arguments.

## Background to the Study of Plato's Psychology

It was natural for Plato to write about the soul in several different ways, as if it had different natures, for the conception of the soul familiar to him from Homer and Hesiod is not singular, but multiple.<sup>9</sup> The archaic and classical Greeks did not have anything like a Formal theory of the nature of the soul; instead, the multiplicity of mental terms in Homer is due not to any conscious division but rather to the variety of experiences which called for their use.<sup>10</sup> As Dodds remarked, the Classical Age inherited a whole series of inconsistent pictures of the “soul” or “self” (179). Appropriately, there is no single meaning for the term *psuchē*, nor is that term the only one commonly used with reference to the soul, personality, character, and emotions of a person. In Homer, Hesiod, the tragedians, and the lyric poets the list includes not only *psuchē* (soul) but *thumos* (spirit), *menos* (spirit or passion), *noos* (intelligence or mind), and *phrenes* (the chest as the seat of passions and the soul).<sup>11</sup>

The term whose usage we should expect to dominate, *psuchē*, is not straightforwardly equivalent to the mind or personality as we now conceive it. In Homer, for instance, *psuchē* is not active when the body is alive nor is it even mentioned in accounts of living persons and their activities. Instead it represents the individual after death and departure for Hades; when the soul of the dead is mentioned, it is generally in terms of *psuchē*. In its visible form, though—the restless shade that haunts graveyards—the soul of the dead is referred to by yet another term, *eidōlon*.<sup>12</sup> *Psuchē* lacks the psychological traits associated with *thumos*, *noos*, and *menos*, and is perhaps better conceived as a breath-soul that keeps the person alive.<sup>13</sup>

In contrast, *thumos* is a person's spirit or “heart,” located in the chest (*phrenes*); it is *thumos* that is the seat of anger and pride, as well as the originator of bold actions. Of the numerous examples of such usages, perhaps the best known occur in the *Medea* and *Iliad* 9. In the famous monologue at *Medea* 1021-80, notably 1078-80, it is her *thumos* that urges her to commit the infanticide: “I know that what I am about to do is bad, but anger (*thumos*) is master of my plans (*boulemata*), which is the source of human beings' greatest troubles.”<sup>14</sup> Similarly, *thumos* in the *Iliad* and the *Odyssey* repeatedly indicates strong emotional response, as it does at *Odyssey* 20.11 where Odysseus' “spirit” (*thumos*) is aroused by the desire to kill his disloyal maids.<sup>15</sup>

The language used with respect to the soul familiar to Plato, then, was varied and reflected several important psychic roles in the living

person: the soul was breath, life, and the seat of violent emotions such as anger, pride, and desire for revenge. In addition to this Homeric understanding of the soul, Plato was very much influenced by Orphic-Pythagorean beliefs which he accepted.<sup>16</sup> While there are differences between the Orphic tenets and those of the Pythagoreans, both strongly emphasized purification (*katharsis*) of the soul. In Orphism *katharsis* was considered necessary because of the soul's imprisonment in the body, an image familiar from Socrates' remarks early in the *Phaedo* (65a, 65b-d, 66b-67b, 67c-d, 81b, 82e, 83d). Both Pythagoreanism and Orphism also accepted a system of moral dualism in which the principles of light, form, and limit were thought to be good in opposition to the principles of dark, formlessness, and the unlimited.<sup>17</sup> Pythagoreans believed in the transmigration of souls into other bodies, a frequent feature of Plato's work that is at odds at times with his other assurances that the soul will eventually end its cycle of reincarnations to exist eternally with the Forms (*Phaedo* 114c4).

To these elements of Homeric and Orphic beliefs about the soul, Plato contributed a new feature: a purely rational or intellectual character which is the essence (*ousia*) of the soul and responsible for its characteristic activities. Given this complex of seemingly contradictory elements, it is no wonder that his psychology is difficult to treat as one coherent theory. There may indeed, for instance, be as many as seven different conceptions of the soul in use in the *Phaedo* alone.<sup>18</sup> These include soul as the element in us that must be cared for if we are to have genuine well-being; soul as the “true self”; soul as intellect or reason, that which apprehends the Forms in a kind of intellectual “vision”; soul as the rational self in opposition to emotion and appetite; soul as the subject of general conscious states; soul in the traditional sense as a life-principle; and soul taken in a general sense to mean a sort of “soul-stuff.” Scholars have long noted the difficulties inherent in sorting out these conceptions and formulating anything like a comprehensible psychological theory in Plato. Still, it is possible to isolate a number of central issues and to shed light on his views in the process of solving the puzzles they present.

## Three Problems in Plato's Psychology

Certain pressing problems motivate Plato's worries and consequently his arguments about the soul: (1) the nature of the soul vis-à-vis Forms, sensible particulars, and alternative models; (2) reconciling if possible the

simple and tripartite models of the soul, as well as understanding what kinds of components or parts constitute it; and (3) the question of immortality. These questions receive their most explicit and detailed answers in the dialogues that form the core of his theory of the soul: the *Phaedo*, *Phaedrus*, *Republic*, *Timaeus*, and *Laws*. First, Socrates works hard in the *Phaedo* to establish the soul's resemblance to the Forms as a basis for its immortality; he also develops three rather involved arguments in the same dialogue against a model of the soul said to be in vogue in some portion of the population: the soul as *harmonia*, or attunement of elements of the body. The second major question concerns the apparent conflict between two ways of presenting the soul, namely the simple soul as pure reason in the *Phaedo* in opposition to the tripartite soul for which Socrates gives a carefully crafted two-stage argument in *Republic* IV. This three-part psychic structure reappears in the well-known and striking chariot metaphor of *Phaedrus* 246a-247c, as does a similar structure in the *Timaeus*, where the three components are physically located in the body (69c-70b). Third, Plato gives us a series of arguments for the soul's immortality: four in the *Phaedo*, and one each in the *Republic* and the *Phaedrus*. How successful are the various strategies pursued in the arguments, and how do they illuminate Plato's conception of the soul?

## 1. The Nature of the Soul

The *Phaedo* gives Socrates the opportunity to argue explicitly for the essential nature of the soul by its analogy with two other entities: Forms and attunements (*harmoniai*). I will examine three lines of argumentation that raise these ontological questions. While the Affinity Argument at 78b-84b is based upon the soul's resemblance to Forms as evidence of its immortality, the Final Argument for immortality at 102a-107a seems, at any rate, to rely upon souls' being something like immanent Forms. In contrast, at the midpoint of the dialogue Simmias presents his thesis that the soul is mortal because it is a kind of attunement (85e-86a), which Socrates considers and rejects in three detailed arguments (92a-95a).

In the Affinity Argument, Socrates first establishes that he and his companions accept two sorts of beings which exhaust their ontology: sensible particulars and Forms. Sensibles have certain properties that entail their destructibility; namely, they are visible, inconstant, and composite. It is just their being composed of parts that entails their dissolu-

bility and hence destructibility. This type of being, Socrates says, is liable to be dissolved by the same parts of which it is constructed; what has been put together and is naturally composite (*suntethenti te kai sunthetô[i] onti phusei*) is likely to break up in the same way it has been put together (78c1-2).<sup>19</sup> In contrast, the beings of which they speak—the beautiful, the good, and the like—are invisible, constant, noncomposite, and thus indissoluble and therefore indestructible (79a-c). Constancy, or the state of being unchanging, is the sure indicator of a simple entity as only the noncomposite has no parts to break off or separate, and hence no mechanism of change. By analogy the soul, which resembles more closely the latter type of being, is at least likely to be indissoluble and thus immortal. Socrates' marked reluctance to state an unconditional acceptance of the proof, indicated in his qualified conclusion, should prevent us from concluding that the soul deserves a place among the Forms; it occupies an uneasy hybrid position between Forms and sensibles, being invisible and perhaps noncomposite yet subject at times to the kind of instability from which sensibles invariably suffer. Earlier seeds of doubts about its final indissolubility had been sown in Socrates' failure to establish a compelling link between constancy and compositeness: the soul is frequently unstable as it is dragged by the body toward objects that are never constant; and it wanders about itself, and is confused and dizzy, as if drunk, in virtue of contact with things of this kind (79c5-9). Only when the soul is near the unvarying objects of its cognition, the Forms, is it unconstant and unvarying, because of its contact with things of a similar kind (79d8). The argument, then, establishes at best a mitigated indissolubility of soul in the weak analogy with Forms, and the soul has features of both sensibles and Forms depending on its cognitive and perceptual activities.

This indeterminate ontological standing is reinvoked in the final argument, about which commentators have debated vigorously concerning the nature and function of the soul. Here the soul has an ambiguous status: depending on how one interprets Socrates' use of "opposites" it may be a substance or "stuff," or an immanent form or character.<sup>20</sup> The argument recalls the principle of opposites Socrates used in the early Cyclical Argument, 70d-71a. Socrates argues that certain things, while not themselves opposites, will not admit one member of a pair of opposites (104c7-9; let us call such pairs of opposites *F* and *G*). These things may be said to bring to whatever they occupy the opposite (*F*) they carry; thus, such things will not admit the other opposite (*G*). The soul, which carries life to whatever it occupies, will thus not admit the opposite of

life, namely, death. Hence, the soul is immortal and imperishable.<sup>21</sup> An example of an entity that acts in the same way is threeness, which, when it is present, carries with it oddness as well. Three will never admit evenness because it necessarily carries oddness. The oddness and evenness present in their respective appropriate numbers are like the large and the small in us mentioned earlier in the dialogue by Socrates (102e5-7). Here he has just distinguished largeness itself (102c6) from the large in us (102e5), which sounds very much like a distinction between transcendent Forms and immanent forms or characters. The soul, bringing life to the body and therefore refusing to admit the opposite, death, may fairly be characterized as an immanent form. There are many difficulties with the text in this argument, over which commentators have argued at length. These include especially what Socrates means when he claims that fire or snow or threeness “occupies” (*katechein*, 104d1, 104d6) or “comes to be” in (*eggignesthai*, 105b9, c3, c9) various sensible particulars. Even the word *athanatos*, usually translated ‘immortal,’ may mean ‘deathless’ or ‘indestructible.’ With respect to the soul the question at hand is both whether or not it must be immortal and also whether it is a particular or else an immanent form.

The second possible psychic model offered is the *harmonia* or attunement of elements of the body, a model that Aristotle also discusses briefly but eventually rejects (*DA* 407b27-408a30). At *Phaedo* 85e-86a Simmias challenges Socrates’ Affinity Argument with the counter-proposal that the soul, even if immaterial, shares the mortal nature of another immaterial entity: the attunement of a lyre, which perishes before the wooden frame and the strings that produced it. Why, Simmias asks, should we not accept that our body is kept in tension, as it were, and held together by hot and cold, dry and wet, and the like, and our soul is a blending and attunement of these same things? (86b7-c1).<sup>22</sup> In this case the soul would clearly perish when the body became unduly relaxed or tight.

What is striking about the *harmonia* thesis is that, while Simmias’ analogy rests upon the patent immateriality of the soul, it has often been taken to be a materialist thesis; i.e., to claim that the soul is composed of *bodily* elements. A material soul would, however, constitute no real threat to the view of immortality Socrates is advancing since he assumes that the soul is immaterial. Only if the *harmonia* thesis is in some sense dualist are Socrates’ three rather involved refutations (92a-95a) of it justified, since it would allow for the existence of an immaterial, but obvious mortal, soul.

These models of the soul are explanatory of the soul as an entire entity without regard to its internal structure. Just as important is the question of the parts of the soul, their nature and number, a compelling Platonic *aporia*.

## 2. Psychic Structure

Perhaps the most celebrated and vexed question of the psychology concerns the structure of the soul. The *Phaedo*’s Affinity Argument, discussed above, depends crucially upon the premise that the simplicity of the soul entails its indestructibility. The simple soul is pure reason, the intelligible entity (80b1) that comprehends the Forms (79a3) and recollects them later (73a1-3). Here in the *Phaedo* the appetites, desires, and emotions are pollutions caused by contact with the body that have not only lured the soul into false beliefs about the nature of the real, but have corrupted the soul so that it will have been interspersed with a corporeal element, ingrained in it by the body’s company and intercourse, through constant association and much training (81c4-6). Such contaminations bind the soul to the body like rivets (83d4), and must be purified by the philosophical life. Conflicts between reason and the emotions and appetites are conflicts between the soul and the body.

This conception of a simple soul essentially identical with the faculty of reason, engaged in the struggle to purify itself of bodily appetites altogether, hardly appears to cohere with *Republic* IV, where Plato offers an extended two-stage argument for the tripartite soul. He introduces immediately a principle about certain opposites that he will apply to the activities and feelings of the soul: “[i]t is obvious that the same thing will not be willing to do or undergo opposites in the same part of itself, in relation to the same thing, at the same time” (436b). So, if a thing appears to do or undergo opposites we may conclude that it is composed of parts. Using this Principle of Opposites, Socrates first argues that the soul demonstrates acceptance and pursuit of a certain object at the same time it demonstrates rejection and pushing away of the same object.<sup>23</sup> Since acceptance and rejection are opposites, two opposite states occur in the soul at once with respect to an object. The example Socrates offers to confirm his judgment is that of drink—but it is, he specifies, thirst for drink *qua* drink, not for, for example, cold drink or good drink. While the person recognizes the drink as an object of appetite, she also understands that she must not drink (as in the case of dropsy), thus pursuing and also

avoiding the drink at the same moment. The presence of these opposite states indicates that it is not with the entire soul that she considers the drink but with two distinct parts: the appetitive component (*to epithumētikon*) and the reasoning component (*to logistikon*). While reason considers the good for the soul as a whole, appetite has concern only for its own apparent good, desiring the drink with no regard for the consequences of consuming it save what it believes will be quenching of thirst.

This concludes the first stage of the argument, which has established two distinct components of the soul. At this point Socrates raises an objection to his own proof: there are actions and states of the soul that seem to originate in neither appetite nor reason. Infants and animals, for example, show a kind of temperament that may be high-spirited or quiet while neither sort of creature possesses reason in the sense of practicing calculation. By the same token, feelings of indignation arise in the soul when one is treated unjustly, feelings not attributable to one's reason or one's appetites. Finally, Socrates discusses the case of Leontius, whose fascination with pale-complexioned boys was apparently well known (439e-440a). Stopping by the site of public executions just outside the city walls, and drawn against his will to gaze on the pale corpses, he is disgusted by his own appetites. The conflict is generated by appetites against *another* psychic force, but this force is not reason (*to logistikon*). Instead, a third part is present in the soul: the spirit or *thumoeides*. This part frequently allies itself with *logistikon* against the desires of *epithumētikon*. Thus, Socrates has argued that the soul is isomorphic with the republic itself, consisting in three components each of which must do its own proper work in the harmonious whole if the soul, and so analogously the republic, is to be just.

In fact, much later in the *Republic* Socrates paints a vivid picture of the very distinctive characters of the three components (588c-e). Here the person's soul is to all outward appearance a single unit; but if one could open the soul to view, one would see three disparate creatures at an uneasy truce within. The appetitive part is a large multicolored and many-headed beast, always changing and capable of growing new and more bestial heads. Spirit is a smaller but fierce lion, and reason is a still smaller human figure, the only completely human component of the soul. Given many attentions and gratifications, appetite enlarges and grows more variegated, crowding out the relatively small reasoning faculty it then easily controls.

There are three noteworthy points in this account of the soul in the *Republic*, taken in light of the previously discussed *Phaedo* passage. First, the soul is now tripartite in some important sense, rather than straightforwardly simple. Second, Plato now assumes that it is the soul's essential nature to vary, not to remain constant. His explanation of psychic states is one of shifting internal alliances in response to external stimuli. Indeed, the educational program of the *Republic* assumes that the soul is tractable, capable of being trained into its proper psychic harmony, with reason ruling and each of the other parts doing its own work. And third, of course, the conflict experienced by the soul has changed from somatopsychic to intrapsychic. It appears that the psychology has undergone a marked change.

Similarly, the parts of the soul presented in the myth of the charioteer in the *Phaedrus* are represented by three components very familiar from the *Republic*. In the *Phaedrus* Plato likens the soul to a chariot with two yoked steeds, driven by a charioteer. It is a union of powers (*dunameis*, 246a): the force of the charioteer guides the soul as reason ruled it in *Republic* IV, while the horses provide the chariot as a whole with its motive power. The noble white horse is *thumoeides*, easily directed and in harmony with the charioteer; but the uncontrollable black horse, rearing and fighting the bit, is unruly *epithumētikon*. Just as in the *Republic*, the soul's progress toward the Forms is the result of resolution of conflict among its three components. Perhaps surprisingly, even the divine souls of the gods are composed of the same three components, although their souls are harmonious because the two matched steeds of spirit and emotion accept absolutely the guidance of the charioteer.<sup>24</sup>

In the discussion above I have used the word "part" freely, but this locution is rather misleading. The terms for part (*meros*, *merē*, *morion*) are introduced in the text only after the argument of *Republic* IV is well begun, at 442b11; and they are not used consistently. Plato often prefers other nouns (as *eidē* [forms] is used at 439e2) or the article with a relative clause (as 439d6 *to men hō[i] logizetai logistikon prosagoreuontes tēs psuchēs*: "we will call the reasoning [thing] in the soul that with which it reasons").<sup>25</sup> This ambiguity in the language used for the components in the soul has led some scholars to conclude that substantive ontological divisions are not established in the *Republic* and *Phaedrus*. Some commentators argue that the soul possesses three functions, aspects, or even motivating forces. Archer-Hind, for example, holds that the soul is reason alone and therefore simple; the division into parts is wholly metaphorical because the three parts of the soul are three differ-

ent modes of the soul's activity under different conditions.<sup>26</sup> Crombie argues that in the *Phaedo*, *Republic*, and *Phaedrus* the soul is reason and that the three parts are merely three aspects of one intellectual love of the Forms,<sup>27</sup> while Guthrie agrees that the soul is essentially simple in all three dialogues, though corrupted by bodily desires.<sup>28</sup> Cornford finds that the three parts are manifestations of a single force, *erôs*, which unifies the soul. Similarly, Grube thinks that the “three parts of the soul are now only three main channels along which [a stream of deep emotion] must flow,” and Taylor refers not to parts but to three “springs of action” in the soul.<sup>29</sup> It is therefore necessary to settle, among other issues, an important mereological concern: if the soul divisions in the *Republic* can be reconciled with the simple soul of the *Phaedo* we need not charge Plato with inconsistency or self-contradiction.

A related question is whether the spirited and appetitive parts of the soul are necessary or contingent to it. It has been suggested by some scholars that the soul has two sorts of natures: it is composite when embodied and noncomposite when disembodied. If so, the lower parts are in some way features of the body added on contingently. Martha Nussbaum argues that the two “lower” parts are contingent in the *Phaedo* but necessary in the *Republic*.<sup>30</sup> Ferrari offers an explanation of his own for the necessity of these components: in the gods’ souls, at least, the spirited and appetitive parts allow the gods to fulfill their responsibility to care for the contingent universe.<sup>31</sup>

The discussion above assumes that the only options for part divisions are a simple soul and a tripartite soul. Plato himself at times, however, seems to divide the soul in two, the rational and the nonrational. Some scholars maintain that Plato uses both the bipartite and tripartite models as needed.<sup>32</sup> For the details of ethical and political theory it was not a sufficient distinction, so the argument goes, whereupon Plato adopted the tripartite model. It may indeed be true that Plato simply used different models of the soul according to requirements of the arguments in particular dialogues.<sup>33</sup> In addition, scholars have frequently noted that the arguments for tripartition may actually do *too much* work, establishing the possibility of an almost limitless number of possible parts to reflect the various psychic conflicts.

Whatever the nature of the components of the soul, there is a further question about possible changes in the soul's constitution from the incarnate to the disincarnate state. If the soul is tripartite, what portion of it goes through the cycle of rebirth? Further, does Plato's theory require that the soul actually be immortal, or could it be sempiternal? Some ar-

gue that the tripartite soul is everlasting, though not immortal, since its creation is detailed in the *Timaeus*’ myth of the demiurge. Many other scholars believe that *psuchê* alone, whether simple or composite, can survive the death of the body.<sup>34</sup> This debate leads us directly to another central point of Plato's psychology: the immortality of the soul.

### 3. Immortality

Establishing immortality was clearly of primary concern to Plato, since there are six arguments for immortality of the soul in the corpus. Four of these occur in the *Phaedo* (the Cyclical Argument, the Recollection Argument, the Argument from Affinity, and the Final Argument). In addition, *Republic* X contains an argument from the specific nature of destructive evils and the *Timaeus* an argument from self-motion. Are these arguments successful, and are they in some way consistent with each other? I consider each briefly in turn.

The first of the *Phaedo*’s arguments, the Cyclical Argument (69e-72d), depends upon the existence of opposites for certain predicates. It is along the following lines:

- (1) If anything *x* comes to be F, and if being F has an opposite being not-F, then *x* comes to be F from being not-F.
- (2) Being alive has an opposite: being dead.
- (3) Whatever is alive has come to be alive. Hence,
- (4) Whatever comes to be alive comes to be alive from being dead.<sup>35</sup>

The argument has often been taken to be extremely implausible, mistaking the nature of predicates ‘alive’ and ‘dead’ in elementary ways. In both cases, the predicates as used in the argument are supposed to stand for attributes that can be acquired by antecedently existing subjects; yet in both cases we think that this is not actually the case. Whatever comes to be alive comes to be existing at that time, so that there is no antecedently existing subject at all, whereas the subject of the predicate ‘dead’ does not in fact continue to exist. In addition, of course, there are also numerous examples of entities such as gold and water that never have been, and are not capable of being, alive or dead. The argument has not, at any rate, been generally taken as a serious attempt by Plato to prove immortality but rather as an initial effort intended to be supplanted by stronger arguments later in the dialogue.

The core of the Recollection Argument (72e-77d), which follows immediately upon the Cyclical Argument, is stated succinctly at 72e3-73b2:

“Yes, and besides, Socrates,” Cebes replied, “there’s also that theory you’re always putting forward, that our learning is actually nothing but recollection; according to that too, if it’s true, what we are now reminded of we must have learned at some former time. But that would be impossible, unless our souls existed somewhere before being born in this human form; so in this way, too, it appears that the soul is something immortal.”

We have, Socrates argues, knowledge of such things as the equal itself. But there is no possible way to acquire such knowledge by sense perception, since the equal is not a sensible object. Hence, our souls must have acquired this knowledge at some time other than during our lives: before birth. Thus our souls existed before their birth into our particular bodies, and we are recollecting, not newly acquiring knowledge of, the Forms.

Because the argument is only half complete (77c1-5), Socrates presents a third argument, the famously flawed Affinity Argument (78b4-84b8). This argument is perhaps the most ambitious of Socrates’ attempts, because it explicitly attempts a proof of the nature of the soul itself, a direct route to immortality. Beginning by postulating two kinds of beings (79a6-7), Socrates establishes three sub-conclusions for similarity of the soul to the Forms: the soul is more similar than body to the invisible, whereas body is more similar to that which is seen (79b16-17); the soul is totally and altogether more similar to what is unvarying than to what is not (79e3-4); and the soul resembles the divine, and the body the mortal (80a9).<sup>36</sup> The major conclusion of the argument occurs at 80b9-10: “soul must be completely indissoluble, or something close (*egus*) to it.”

The last of the *Phaedo*’s arguments for immortality occurs near the end of the dialogue, catching up and weaving together the themes of causation, opposites, and immortality to claim that the soul, which comes to the body bringing life, will never admit death but will depart the body instead. Plato has been accused here in the scholarly literature of committing the fallacies of equivocation, composition, and *ignoratio elenchi*, having proven only, it is claimed, that the soul is still a soul so long as it is alive.<sup>37</sup> If this charge is true, then what is considered the most successful of the arguments from immortality in the *Phaedo* will be irreparably flawed.

An entirely different strategy underlies the argument for immortality at *Phaedrus* 245c5-246a2. Here the claim is that soul is immortal not because it is completely constant and therefore incomposite, but because it is self-moving.<sup>38</sup> The argument owes a debt to Alcmaeon, who argued that what moves unceasingly is immortal, although he importantly did not distinguish between the self-mover and that which is moved by something else.<sup>39</sup> The argument opens ambiguously: *psuchē pasa athanatos* (all soul is immortal). Here the lack of initial article makes possible two readings, on which scholars have disagreed: ‘all soul,’ soul taken collectively in whatever Form it occurs; or ‘every soul,’ the distributive sense of individual souls taken as a group.<sup>40</sup> The sense which is important for the dialogue is predominately the former, since the individual souls in the metaphor of the chariot are presumed to survive death and to climb, with varying degrees of success, to the Forms (247c-248e). It is as individual souls, bearing the burdens of undisciplined incarnate choices or philosophically purified, that they are reincarnated into the appropriate bodies (248e-249d).

The series of four arguments itself establishes that (1) only what is self-moving (*to hauto kinoun*) never ceases to move, as it cannot abandon its own nature (245c1-8); (2) the self-mover is a first principle (*archē*) and thus can neither come into being nor go out of being (d1-6); (3) the self-mover and the first principle of motion are the same, and as such are immortal (d6-e2); and (4) finally, to say that the self-mover is the soul is to give the essence and account of the soul (*psuchēs ousian te kai logon*, e2-246a2). In contrast to the *Phaedo* arguments for immortality, all of which rely upon the soul’s resemblance to the unchanging Forms, the argument of the *Phaedrus* introduces a conception of the soul that would have been intuitively obvious to their audience: the soul as principle of movement, what distinguishes between the living being and artifacts or other natural but inanimate objects incapable of attaining motion on their own. The contradictory metaphysical needs of both sets of arguments should, however, signal the quandary in which Plato finds himself as he tries to combine features of the soul required by his metaphysical theory with features so common as to be unarticulated.

The *Republic* also contains an argument for immortality at 608c-611a, although the argument has been held rather in disrepute by many commentators—to the point of being cited as one of the few really embarrassingly bad arguments in Plato.<sup>41</sup> I sketch the rough framework as follows:

- (1) There is a natural good and bad for everything.
- (2) When a thing's natural evil attaches itself to that thing, it corrupts and then destroys it.
- (3) If that natural evil does not destroy it, then nothing else will do so.
- (4) The natural evil of the soul is vice: injustice and the other vices.
- (5) Vice does not, however, destroy the soul. Hence,
- (6) There is nothing that will destroy the soul. Hence,
- (7) The soul, as indestructible, is also immortal.

Critics have raised numerous strong objections to this argument, while others have tried to repair its weaknesses.<sup>42</sup> Three problematic features frequently targeted by scholars are the singular natural vice claim, the premise that vice does not destroy the soul, and certain background assumptions of Plato's metaphysics (e.g., the dualism of body and soul).<sup>43</sup> Given such obvious counterexamples as healthy wood, paper, and even human bodies that can be destroyed by fire as well as by rot, insects, or disease, it seems absurd to claim that there is only one natural evil for any given object. About premise (3) we might ask why it is not possible that the soul might survive several incarnations incrementally weakened by its accumulated vices, to perish finally at some final disembodiment. And last, any theory of the mental that rejects dualism—such as mind-body identity theory or supervenience theory—will have ready responses to the view that the soul could at any rate survive the death of the body. In short, although Plato had good reasons to reject a strategy of arguing for the soul's immortality on the basis of its strict simplicity, given *Republic* IV's tripartition argument, still the strategy he chose has certain undeniable drawbacks.

The *Timaeus* contains yet another account of the soul, which is composed of elements assembled by the Demiurge in the case of the World Soul and of the dregs of the same mixture by the lesser gods in the case of human souls. While the human soul is composed of reason, spirit, and appetites, only the reason is created by the Demiurge, while the two lower parts are added to it from mortal material (41d-44c). The soul possesses only one immortal part, the reason, which is established in the head of the mortal body. The other two parts, situated in different parts of the body, are built into that body as it is fashioned: spirit (*thumos*) is in the chest, and appetite (*epithumētikon*) appropriately situated in the belly (69d-71a). These components are mortal, as they are constructed by the lesser deities from the same matter of which the body is made (69c7-8).

While Plato thus assumes that the lower parts do not survive the body's death and so do not return to the star homes from which the immortal souls originate, nowhere does he account for their reappearance with a new body upon reincarnation. The natural question is whether these parts are supposed to be generated each time from the body itself or whether they are somehow inert as they accompany the reason between incarnations.

The importance of settling the mereological issues lies in the possible threat to the simplicity of the soul, and hence to its immortality. Plato has argued in the *Phaedo* that the soul is at least likely to be immortal specifically because it is simple. If the soul has parts which truly render it composite, then its immortality is threatened, or at any rate, one of the central arguments for immortality is threatened. Thus, Plato may have to make an unhappy choice. On the one hand, a tripartite psychology explains phenomena of internal conflict which he has observed; it also allows him to postulate psychic justice that is isomorphic with justice on the larger scale of the ideal republic. This tripartite conception, however, will at the least achieve immortality only with difficulty, as it is not easy for anything composed of many things to be immortal if it is not put together in the finest way (*Republic* 611b).<sup>44</sup> On the other hand, if the soul is noncomposite, its immortality is well-nigh guaranteed. The importance to Plato of the soul's immortality is clear from the number of different arguments he adduces in the proof of it. He must, then, settle the issues of the nature of the soul and in turn of parts it may have; the soul is a crucial component of his metaphysics and necessary to his account of knowledge.

In what is commonly taken to be the last of Plato's works, the *Laws*, he does seem to resolve at least one of the difficulties discussed above: the agentlike parts of the *Republic* are replaced with singular agents who experience the various emotional and cognitive states belonging to them. The passage at 892a-899d establishes the soul as that which is defined as self-moving motion (896a1-8), and its motions are listed at 897a1-4: wish, reflection, forethought, counsel, opinion true and false, joy, grief, confidence, fear, hate, love, and all similar motions. Here the central point of the argument for immortality at *Phaedrus* 245c-246a returns, as the Athenian gives responsibility for all motion to soul, which causes in turn every sort of movement in the universe. Neither the static soul of the *Phaedo* nor the composite soul of the *Republic* is retained in Plato's final treatment of the soul.

## This Volume

This brief overview of the psychology has sketched the relevant arguments from Plato's dialogues. I now turn to the articles in this collection, which investigate three major problems discussed above: the nature of the soul (I), the tripartite soul (II), and arguments for the immortality of the soul (III). To start, the first two articles discuss theses about the nature of the soul; first, whether in the *Phaedo* it is an immanent Form, and secondly, whether it might be a *harmonia* or attunement. Weller focuses on the problems of the Final Argument for immortality at *Phaedo* 102a10-107b10, which has been disdained by a number of scholars as obviously fallacious. The argument appears to establish not that the soul is immortal, but only that it is a soul so long as it is alive. In addition, the argument appears to rely upon the assumption that the individual soul is a sort of immanent Form, although there is no Form of Soul which it would instantiate. Weller argues that the soul is not an immanent Form, and that the soul brings life to the body because of its causal powers, which are grounded in the soul's intrinsic properties. The Final Argument is not fallacious (as charged by Keyt) because the soul is alive, not in the way that tallness is the-tall-in-Socrates, but in the same way in which ordinary snow is cold and cannot admit being hot.

Taylor considers Simmias' presentation at *Phaedo* 85e-86d of the thesis that the soul is a kind of *harmonia* or attunement of bodily elements. After clarifying the various possible meanings of the term *harmonia*, he goes on to discuss all three counterarguments of Socrates at 92a6-95a2. Taylor's analysis of the last two arguments, 92e4-95a2, exposes the fallacy in the section of text he terms Argument B, and gives a new explanation for what has gone wrong. If Plato took himself to have provided arguments that definitively refuted the *harmonia* thesis, his development of the tripartite model of the soul in the *Republic* and the *Phaedrus* would be unmotivated.

In "Supervenience and the Thesis That the Soul Is a *Harmonia*," Wagner offers a new interpretation of the thesis that the soul is a *harmonia*. It is, the author argues, a supervenient dualist view that constitutes a strong threat to Plato's own substance dualism. Commentators have reviewed this thesis in some detail, as they have also analyzed and assessed Socrates' counterarguments at 92a-95a. What they have *not* done, however, is to explain why Socrates considers this thesis so powerful that he devises three quite sophisticated arguments against it. Because the *harmonia* is separate from the body it meets an important criterion for being

a substance; thus, it is an immaterial substance causally dependent upon the material body and hence necessarily mortal. On this reading, Simmias' counterexample of the *harmonia* is indeed a threat to the immortality of the soul.

In Section II we turn to considerations and criticisms of this tripartite soul. Cooper argues that properly understood, Plato's theory is both subtle and accurate in its account of the psychology of human motivation. He suggests that it is the psychological theory, in fact, from which Plato derived the political analogy of the ideal republic's three classes, rather than the other way around. Analysis of the three components of the soul presented in *Republic* IV shows what is distinctive about each, and hence why there are precisely three parts of the soul in Socrates' two-stage argument. On Cooper's view, the parts are independent sources of motivation: reason provides motivation for knowing the truth and for ruling; appetite contains desires originating in facts about our bodily constitution, our experiences, and the pleasures that result from such experiences; and spirit is united by competitiveness, desire for esteem, and self-esteem.

Smith raises anew questions about the analogy between the soul and the state in the *Republic*. He argues that Plato's arguments for tripartition of the soul contain an equivocation, which prevents him from validly proving that the soul has three parts. This does not prevent Plato from developing an account of justice, however, even if the tripartition argument fails. Because justice in a city or individual consists in doing one's own *pros ti* (with respect to something), it is doing one's own with respect to the proper functions of the soul that constitutes justice in a soul, and this conception is not dependent upon tripartition.

Shields, in contrast, argues that the standard developmental account of Plato's move from the simple soul of the *Phaedo* to the tripartite soul of *Republic* IV is wrong. Neither the *Phaedo* Affinity Argument nor the picture of the soul in *Republic* X is inconsistent with the argument for tripartition because the latter fails to establish that the parts of the soul are *necessary*. Shields develops a new mereological account on which the "part-generating principle" of *Republic* IV is strictly empirical.

Williams considers another problem with the analogy between the soul and the state: the analogy contains an ineliminable tension. A city, to be just, is required to have a reasoning, a spirited, and an appetitive element in it, exactly as a soul must; but then the individual citizens representative of the elements are not just. Hence, the city does not have a majority of men who are just, which contradicts Socrates' statement that

a city is just if and only if its men are. In response to Williams and other critics, Lear argues that Plato's analogy of state and soul is not fallacious. Further, there is no such tension in the account of the sort that Williams discovers. To understand correctly how the analogy works, we need to see that it is a by-product of a certain psychological dynamic consisting in internalizations and externalizations across the boundary that separates state and individual soul. In fact, together the state and soul constitute a whole psychological universe. Applying this insight to Plato's critique of the poets, Lear suggests that once we understand Plato's psychology we can see that it is the internalization of the poets' words in the soul that is problematic.

The second section of the collection closes with a discussion of important points concerning the tripartite soul: problems arising from Plato's characterization of the three components of the soul as agentlike. Bobonich turns to questions about Plato's conception of the parts in the soul in the *Republic* and the *Laws*, discussing *akrasia* in the context of theories of the soul presented in these dialogues. In a careful examination of the theory of the parts in the *Republic*, which are agentlike (the Partition Theory), he notes the unsatisfactoriness of this psychological model, which causes certain problems. These include the plausibility of the partitioning, the generation of too many agents in the soul, and the unity of the self. The late psychology of the *Laws*, which omits such agentlike parts in favor of agents and their affections—namely, occurrent beliefs, desires, and emotions—is a more successful theory of akratic action than that of the *Republic*.

In Section III the authors examine various arguments for the immortality of the soul, which Plato needs to establish as a buttress for his theory of recollection. Bostock provides a guide through the difficult passages regarding the soul in the *Phaedo*. He discusses the meanings of the term *psuchê*, the soul's activities, the disembodied soul, implications for morality, the relation between soul and person, and the set of four arguments for immortality. Moving to criticism of a particular argument for immortality in the *Phaedo*, Gallop raises several issues about the importance and viability of the Cyclical Argument at *Phaedo* 69e6-72e2. He argues that the argument is intended to prove immortality of the soul as well as that it is intended to complement the Recollection Argument: together they form a unit which covers both the existence of the soul before birth (Cyclical Argument) and also its existence after the death of the body (Recollection Argument). Gallop points out that the Cyclical Argument depends crucially upon a principle of opposites by which be-

ing alive and being dead are exhaustive of the options for beings. There are two difficulties with this principle, however: alive and dead are not exhaustive, since there are entities that are neither; and it is not true that opposites in the sense used in the argument always come from each other. In the end, the Cyclical Argument is a self-contained argument Plato raises and then finds defective on the grounds that it does not prove the *post mortem* existence of every soul in every one of its incarnations. For that, Plato provides the Final Argument (102a-107a). Frede takes up the defense of the latter proof, arguing against its critics that it is valid, not invalid, and that Plato does not beg the question by assuming that the soul is alive. She defends the argument's inferences and points out certain assumptions Plato may have made. Ultimately, however, the argument is flawed because it fails to define the nature of the soul, which violates the Socratic rule that one must properly grasp the nature of a thing before arguing that that entity does or does not possess a certain property.

In contrast to the *Phaedo*, the *Republic* contains only one argument for immortality, a proof that has been roundly criticized by scholars. Brown's rehabilitative efforts take us methodically through the argument, which is based upon each thing's having its own singular natural evil. Rather than dismissing the argument as unsalvageable, Brown points out its strengths. Critics have attacked three points in the argument: the essential destructibility claim that not even evils can destroy a thing that lacks a natural evil, Glaucon's statement that vice does not ever destroy the soul, and Plato's assumptions of soul-body dualism and immortality of identifiably individual souls. Understood within the context of the *Republic* as a whole, the argument is not so weak as commentators have often proclaimed it to be.

The final two articles in this section concern the argument for immortality at *Phaedrus* 245c-246a. Robinson carefully dissects the argument at *Phaedrus* 245c-246a into four subarguments, laying bare its structure. He points out that the proof solves difficulties of the psychology presented in the *Phaedo*, where the soul is static, by acknowledging the soul as the source of motion. Finally, he relates the conception of the soul as self-mover in the *Phaedrus* to problems with the precosmic chaos discussed at *Timaeus* 49a ff. It is the soul as self-mover that best captures the soul's function as life principle. In the last article, Bett takes us carefully through the same proof to clarify the place of the *Phaedrus* in Plato's works, as well as the role of the argument in the dialogue as a whole. He investigates the argument in detail, relating it to the famous

myth of the soul as chariot with charioteer and horses as well as to treatments of the soul in several dialogues: the *Phaedo*, *Republic*, *Sophist*, *Timaeus*, and *Laws*. He argues that the myth of the charioteer and horses shares assumptions with the argument for immortality about the nature of the soul, especially that the soul is eternally in motion. The *Phaedrus* view of the soul, he concludes, is closer to those put forward in the *Laws*, the *Sophist*, and the *Timaeus* than it is to those of the *Phaedo* and the *Republic*.

Following the strands of argumentation through the dialogues, what the student of Plato's psychology might find fascinating above all are his herculean efforts to create from a collection of fragmented conceptions of the soul and the metaphysical needs of the theory of Forms a reasonably consistent view of the soul, one that will be able to account for knowledge and moral reasoning in human beings. Given a conglomerate of Homeric ideas about the breath-soul that is also the seat of the emotions; the traditional picture of it as a shade that suffers an attenuated existence in Hades; and views about transmigration, immortality, ritual, and purification from the Orphic-Pythagorean cults, Plato adds to the mix his own conviction that reason above all is essential to the soul. He never distinguishes an absolute substantive difference between body and soul, or person and soul, yet his attempts in both *muthos* (myth) and *logos* (argument) to define and understand the soul remain intellectually gripping.<sup>45</sup> The arguments for immortality show a breadth and ingenuity that ought to command our respect for sheer inventiveness and dogged determination. Because the philosophical project of purification by the life of reason presumes properties of the soul that will prevent it from being placed absolutely in the category of Forms, the very life Plato advocates so passionately in the figure of Socrates throughout the dialogues is made both necessary and impossible by the limitations of the soul.<sup>46</sup>

### Appendix: Problems in Plato's Chronology

In order to study any area of Plato's philosophy, scholars must first solve a general problem with his works: neither Plato himself nor any of his contemporaries has given us the precise dates of the dialogues' composition. If this chronology is not determined, it is difficult or impossible to trace possible changes in Plato's views in any given area. This problem has not gone unnoticed by commentators, many of whom have been engaged in a lively debate that has resulted in a prodigious amount of

scholarship over the past four decades. Still, the current debate about the correct way to date and then interpret the Platonic corpus is certainly not new; indeed, it is only a reworking of complications and conflicts in earlier frameworks that were developed beginning almost immediately after Plato's death, in the Academy by his successors.<sup>47</sup> Beginning with Speusippus and Xenocrates, continuing with Arcesilaus and later the Neoplatonists (especially Plotinus), through the nineteenth century, commentators generally attributed to Plato a body of doctrines presented quite consistently throughout the corpus. Plato scholars were unitarians, who read the dialogues in a number of different orders and treated them as presenting a single unitary conception of any particular doctrine.<sup>48</sup> Nineteenth- and twentieth-century scholars, on the other hand, seriously debated the issue of unity and development in the dialogues' various doctrines—if indeed Plato's philosophical positions can correctly be called doctrines. Modern commentators, following in the wake of Campbell's landmark study of 1867, have been attempting to unravel the mystery of the dialogues' respective dates and to establish a definite chronology for them using a number of varied, sophisticated criteria.<sup>49</sup> Linguists use stylometric methods to analyze texts, counting occurrences of certain features of Plato's style.<sup>50</sup> Such linguists count occurrences of characteristic particles, measure changes in the avoidance of hiatus, and note differences in *clausula*—the rhythms of prose line endings—in which Plato's usage follows distinct patterns. Several other kinds of evidence are adduced in arguments about the chronology of the dialogues, as well. For instance, there are occasional internal cross references to other dialogues; two such examples are the conclusion of the *Theaetetus*, which glances forward at further conversations with the same characters that will occur in the *Sophist*, and the opening remarks of Timaeus in the *Critias*, upon having completed the long march (106a2) of his previous argument in the eponymous dialogue. References to external events also mark the temporal limits of certain dialogues: notably, the *Theaetetus* opens with reference to the battle in which Theaetetus himself was wounded and is now dying of wounds and dysentery, in 369 B.C.E. Our final sources of information about chronology are the scattered remarks of various ancient commentators.<sup>51</sup>

Scholars have at times tried another strategy as well, arguing that it is possible to date the dialogues according to what seem to be natural developments in doctrines or arguments used by Plato.<sup>52</sup> This strategy has proven unsuccessful at best and circular at worst. Unfortunately, the cumulative results of all of these kinds of evidence, even including the very

painstaking and exhaustive stylometric analyses done with the use of computers, are inconclusive. This is partially due to the marked rarity of external references and internal dialogic markers. In the case of stylometry the inconclusiveness is the result of several factors, especially the difficulty of controlling for various factors that govern special cases of Plato's linguistic usage, including the genre, his subject matter, and his purpose (e.g., didactic, rhetorical, etc.); and the possibility that Plato revised his work, incorporating later stylistic features into earlier works.<sup>53</sup> It is also nearly impossible to extricate the researcher's subjective judgment from the selection factors and decisions about, for instance, the length of prose units to be counted and from where in the dialogues they ought to be culled. Given all of these factors, the picture for arriving at a reasonably secure chronology in Plato looks very bleak indeed.

Recently, commentators have acknowledged this fact. Julia Annas briefly discusses the relevance of the studies above to the possibility of establishing anything like a development of Platonic doctrines in the dialogues, concluding that all hopes of establishing such a scientific ordering failed to materialize (26). The work of Debra Nails also establishes grounds for deep skepticism about all of the various chronologies proposed.<sup>54</sup> In the recent Hackett edition of the complete dialogues, editor John Cooper urges readers not to undertake the study of Plato's works holding in mind the customary chronological groupings of "early," "middle," and "late" dialogues. It is safe, he goes on to say, to recognize only the group of six late dialogues.<sup>55</sup> These six late dialogues now generally accepted by Plato scholars, and also established by Brandwood, are the *Timaeus*, *Critias*, *Sophist*, *Statesman*, *Philebus*, and *Laws*. Beyond the limits of this last group, no more determinate set of groupings is universally accepted, and even within groups, for those who argue for them, no order has been determined.<sup>56</sup>

Still, some scholars accept the possibility of chronology, and for them the ordering of the dialogues generates hypotheses about development of Plato's thought, including his views of the soul. Gregory Vlastos, who practically singlehandedly revived the field of Plato studies in the 1960s, argued strongly for definite clusters of early, middle, transitional, and late dialogues, as well as for a defensible order within some of these groups.<sup>57</sup> Other philosophers have argued for similar dialogue groupings, often agreeing on at least one basic cluster—for example, that the early dialogues are elenctic or Socratic works that end aporetically after an extended *elenchos* with at least one unruly, but ultimately ineffectual, interlocutor.<sup>58</sup>

The significance of rejecting such attempts to arrive at an order of composition is, of course, that one denies the possibility of tracing views or doctrines systematically through the corpus in a chronological development of Plato's thought. We should not despair, however, of studying Plato's works with philosophical integrity; it is still possible to trace the *logical* development of his ideas about the topics he explored.<sup>59</sup> For this reason, this volume treats the issues under consideration as questions Plato answered in ways that differ according to dialogic aims, and does not assume any determinate chronological ordering.<sup>60</sup> While certain arguments present *prima facie* contradictions, these need neither be accepted as ineluctable nor attributed to any chronological development of Plato's ideas. This is particularly important for the psychology, the bulk of which is laid out in the second group of dialogues, where arguments for ordering them are tenuous and deeply contested. Despite the ultimate failure to establish the order of composition of Plato's dialogues, however, there are approaches to the study of Plato, in particular to the study of his psychology, that are fruitful and interpretatively sound.

## Notes

1. There have, however, been several fine monographs on the subject, including especially T. M. Robinson *Plato's Psychology* (Toronto, Ontario: University of Toronto Press, 1995); A.-Ed. Chaignet, *De La Psychologie de Platon* (Bruxelles: Culture et Civilisation, 1966), and Léon Robin, *La théorie platonicienne de l'amour* (Paris: Presses universitaires de France, 1964).
2. C. D. C. Reeve, *Philosopher-Kings: The Argument of Plato's Republic* (Princeton, N.J.: Princeton University Press, 1988), 169.
3. See Jerry Fodor, *The Modularity of Mind* (Cambridge, Mass.: MIT Press, 1983), "The modularity of mind," in *Meaning and Cognitive Structure*, ed. Z. Pylyshyn (Norwood, N.J.: Ablex, 1986); "Why should the mind be modular?" in *Reflections on Chomsky*, ed. A. George (Oxford: Blackwell, 1989), reprinted in J. Fodor, *A Theory of Content and Other Essays* (Cambridge, Mass.: MIT Press, 1990). Other accounts and responses to Fodor include J. Garfield, ed., *Modularity in Knowledge Representation and Natural-Language Understanding* (Cambridge, Mass.: MIT Press, 1987), T. C. Meyering, "Fodor's modularity: A new name for an old dilemma," *Philosophical Psychology* 7 (1994): 39-62, and E. Olsson, "Coherence and the modularity of mind," *Australasian Journal of Philosophy* 75 (1997): 404-11.
4. The literature on functionalist views of mind is voluminous. For a general introduction see Ned Block, ed., *Readings in Philosophy of Psychology* (Cam-

bridge, Mass.: Harvard University Press, 1980), Part Three; also David Rosenthal, ed., *The Nature of Mind* (Oxford: Oxford University Press, 1991), III.B.

5. Here see especially the seminal work of David Chalmers, who formulated this version of the problem of consciousness in *The Conscious Mind* (Oxford: Oxford University Press, 1996), and Ned Block, Owen Flanagan, and Güven Güzeldere, eds., *The Nature of Consciousness* (Cambridge, Mass.: MIT Press, 1997). The new *Journal of Consciousness Studies*, the Formation of the Association for the Scientific Study of Consciousness, and the annual Tucson conference on consciousness at the University of Arizona, in addition to many scholarly and scientific books and articles both in print and electronic form, attest to the enormous growth of interest in topics at the nexus of philosophy of mind, neurophysiology, and psychology. Recent discussions in print and on e-mail philosophy of mind and cognitive studies lists, while wide-ranging, have centered on consciousness as a cluster concept rather than a singular phenomenon.

6. The most recent attempts to come to grips with this problem are, notably, Richard Swinburne, *The Evolution of the Soul*, 2d. ed. (Oxford: Oxford University Press, 1999) and Colin McGinn, *the Problem of Consciousness* (Oxford: Blackwell, 1991) and *The Mysterious Flame: Conscious Minds in a Material World* (New York: Basic Books, 1999). McGinn argues provocatively in both works that the human intellect can never know consciousness, which necessarily challenges Plato's investigative framework from its inception.

7. F. M. Cornford, *The Republic of Plato* (Oxford: Oxford University Press, 1945), xxvii.

8. There are also general problems in any study of Plato's philosophy, founded largely upon the difficulties of determining an accurate chronology of the dialogues. For a more technical discussion of these issues, interested readers should turn to the Appendix.

9. Jan Bremmer, *The Early Greek Conception of the Soul* (Princeton, N.J.: Princeton University Press, 1983), 66. Bremmer's broad study of ancient conceptions of the soul draws also upon discussions of views of contemporaneous cultures of central Asia. The standard references for such studies, though now amended with respect to Homeric conceptions, are: E. R. Dodds, *The Greeks and the Irrational* (Berkeley: University of California Press, 1971); R. B. Onians, *The Origins of European Thought*, 2nd. ed. (Cambridge: Cambridge University Press, 1954); Erwin Rohde, *Psyche: The Cult of Souls and Belief in Immortality among the Greeks*, trans. W. B. Hillis, 8th ed. (London: Routledge & Kegan Paul, 1950); and the classic although contested work of Bruno Snell, *The Discovery of the Mind* (New York: Dover, 1982).

10. E. L. Harrison, "Notes on Homeric Psychology," *Phoenix* 14 (1960), 67.

11. The last of these terms, *phrēn* or *phrenes*, is the diaphragm or chest, where frequently *thumos* and *menos* are located, and even at times the *noos* (later *noûs*).

12. See Bremmer, *Early Greek Concept*, 16.

13. Shirley M. Darcus, "A Person's Relation to Ψυχή in Homer, Hesiod, and the Greek Lyric Poets," *Glotta* 57 (1979), 34. See the continuation of this article by Darcus in the same volume of *Glotta*, 167-73. Also of interest are K. Von Fritz, "Noos and Noein in the Homeric Poems," *Classical Philology* 38 (1943): 79-93 and "Nous, Noein, and Their Derivatives in Pre-Socratic Philosophy (Excluding Anaxagoras)," Part II (*Classical Philology* 41 [1946]): 12-34.

14. This translation is that of Christopher Gill, *Personality in Greek Epic, Tragedy, and Philosophy* (Oxford: Clarendon, 1996), 223. See also his discussion of the same passage on 217, where he argues that Medea is distancing herself in the speech from her *thumos* at the same time that she reaffirms her plans.

15. Throughout *Iliad* 9 there are uses of *thumos* as emotional; see 255, 260, 496, 595, 629, 635. Gill's arguments against Snell in chapter 3 of *Personality in Greek Epic, Tragedy, and Philosophy* are thorough and convincing.

16. There is no systematic way to distinguish the Orphic beliefs and practices from those of the Pythagoreans; hence my reference to that collective body of cult views. W. K. C. Guthrie has discussed these questions thoroughly in *Orpheus and Greek Religion* (Princeton, N.J.: Princeton University Press, 1993), especially chapter VII.

17. See Guthrie's discussion at *Orpheus and Greek Religion*, 219.

18. T. M. Robinson, *Plato's Psychology*, chapter 2, surveys these meanings, as does Gallop in his commentary, 88-91. The list of seven views of the soul is drawn from Gallop.

19. Translations of the *Phaedo* are based on those of Gallop in his commentary.

20. Gallop's commentary analyzes the arguments at length.

21. Gallop's commentary on this argument presents two alternative versions of the argument based on two possible translations. Version A involves, as does my summary in the text, pairs of opposites in "stuffs" or objects; version B makes use of pairs of opposite Forms. If one adopts the translation and corresponding version B, one is committed to a Form of Soul and souls as instances of that Form. See Gallop, 199-222.

22. Simmias introduces the theory as what we actually take the soul to be (86b7). The referent of 'we' in this line is unclear; presumably Simmias means the Pythagoreans, yet the *harmonia* thesis is inconsistent with Pythagorean teachings on suicide (and on transmigration). Gallop's commentary, 147-49, is helpful here, as is H. B. Gottschalk, "Soul as *Harmonia*," *Phronesis* 16 (1971): 179-98.

23. Michael Woods in "Plato's Division of the Soul," *Proceedings of the British Academy* 73 (1987): 23-48 adopts this term from previous commentators; Irwin in *Plato's Ethics* dubs it the Principle of Contraries (204), noting that it certainly does not apply to contradictions; and other scholars use similar terminology. The important question here is whether it is a logical or an empirical principle. Some, notably I. M. Crombie in *An Examination of Plato's Doctrines*, 2 Vols. (New York: Humanities Press, 1963) and W. K. C. Guthrie in *A History*

of Greek Philosophy, Vol. IV (Cambridge: Cambridge University Press, 1987), argue that it is an instance of the Law of Non-Contradiction and thus purely logical. Shields, "Simple Souls" in *A Union of Powers: New Essays on Plato's Psychology* (forthcoming) [this volume] defends the position that the "part-generating principle" is empirical, derived as it is from Socrates' experience of human psychology, with the result that the argument for tripartition cannot produce necessary parts of the soul.

24. Clearly these are not the Olympian gods in their full idiosyncratic character, although Socrates does later advert to personality types who characteristically worship one or another of Zeus, Apollo, etc.

25. A number of translators of the *Republic* add part language where it is not in the Greek, creating the illusion that Plato had in mind the kind of distinct ontological divisions which are the very thing at issue here. Reeve's admirable translation of the work is guilty of this practice.

26. Archer-Hind, *The Phaedo of Plato* (New York: Arno Press, 1973), xxxiv.

27. See Crombie's discussion Vol. I, 343-59.

28. Guthrie Vol. IV, 421-25.

29. F. M. Cornford, *The Unwritten Philosophy*, (Cambridge: Cambridge University Press, 1950), 71; G. M. A. Grube in *Plato's Thought* (Boston, Mass.: Beacon Hill, 1958) 136 claims that the division of the soul into three parts is a very great advance (133) over the simple soul but does not explore what kinds of parts this division creates. Also see A. E. Taylor, *Plato: The Man and His Work*, (New York: Meridian Books, 1956), 281-82.

30. Martha C. Nussbaum, *The Fragility of Goodness* (Cambridge: Cambridge University Press, 1986), chapter 7.

31. G. R. F. Ferrari, *Listening to the Cicadas* (Cambridge: Cambridge University Press, 1987).

32. See, for example, Paul Shorey, *The Unity of Plato's Thought* (New York: Archon Books, 1968), 42-43.

33. T. M. Robinson, *Plato's Psychology*, chapter 11, neatly sums up the different pictures of the soul evoked for Plato's various purposes in the dialogues.

34. James V. Robinson, "The Tripartite Soul in the *Timaeus*," *Phronesis* 35 (1990): 103-10, argues for the tripartite immortal soul, while Christopher Rowe in *Plato: Phaedo* (Cambridge: Cambridge University Press, 1993), Hackforth, Guthrie, Bett, and T. M. Robinson oppose this view.

35. This argument is taken in modified form from Bostock's commentary on the *Phaedo*, 51.

36. The structure of the argument is disputable. David Apolloni, "Plato's Affinity Argument for the Immortality of the Soul," *Journal of the History of Philosophy* 34 (1996): 5-32, notes the history of the three-part structure I described above, which is the obvious surface division of the text. Yet it also appears that Socrates is arguing for a set of entailments: what is invisible is con-

stant, what is constant is necessarily noncomposite, what is noncomposite is indissoluble because it has no parts to be broken off or lost, and thence to the immortality of the indissoluble. This line of entailments explains the list of attributes appearing at 80b1-6, many of which have not been included in the previous argumentation.

37. See especially David Keyt, "The Fallacies in *Phaedo* 102a-107b," *Phronesis* 8 (1963): 167-72, for detailed arguments for these claims. Weller, "Fallacies in the *Phaedo* Again," *Archiv für Geschichte der Philosophie* 77 (1995): 121-34, [this volume] responds. See also David Gallop, *Plato: Phaedo*, 203-7 for two versions of the argument derived from alternate translations.

38. The quite recent discovery of an Oxyrhynchus papyrus (1017) makes available this reading (*autokinētos* or self-moving instead of the previously accepted *aeikinētos* or always moving). On this, see T. M. Robinson, *Plato's Psychology*, 111 and his thorough discussion in chapter 6. Charles Griswold, *Self-Knowledge in Plato's Phaedrus* (University Park: Penn State Press, 1986), 78-87, accepts the same reading, as does Ferrari in *Listening to the Cicadas* (Cambridge: Cambridge University Press, 1987), 123-25. Hackforth, on the other hand, rejects the reading of Oxy. 1017 by Robin and others, on the grounds that it is defeated by the logic of the passage as a whole. See Hackforth, *Plato's Phaedrus* (Cambridge: Cambridge University Press, 1985), 65.

39. Jonathan Barnes in *The Presocratic Philosophers* (London: Routledge, 1989), 115-17 discusses the relevant fragment (DK A12) of Alcmaeon and commentators on that fragment, including Aristotle (*de Anima* 405a29-b1).

40. See Hackforth's commentary 63-68.

41. The judgment belongs to Annas, *An Introduction to Plato's Republic*, 344. See Brown, "A Defense of Plato's Argument for the Immortality of the Soul at *Republic* X 608c-611a," *Apeiron* 30 (1997): 211-38 [this volume] for a recital of detractors' remarks.

42. While Annas disparages the argument, Reeve in *Philosopher-Kings* leaves serious doubts unanswered at the close of his discussion (160); Grube finds it an awkward argument (138). White in *A Companion to Plato's Republic* (Indianapolis: Hackett, 1979), 259-60 calls it a "far-from-cogent" proof.

43. In section II of "A Defense of Plato's Argument for the Immortality of the Soul at *Republic* X 608c-611a," Brown very clearly lays out and responds to six criticisms of the argument derived from these three sorts.

44. The passage is translated by Grube, Reeve, Shorey, and others so that it attributes many parts to the soul. The Greek is, however, just *suntheton te ek pollōn*, or composed of many unspecified entities.

45. One should not believe, however, that there is a strict line that divides myths, or fictional accounts, from arguments. Plato at times uses the terms nearly interchangeably.

46. I am grateful to Christopher Shields for judicious, helpful comments on earlier drafts of this essay. All errors are, of course, my own.

47. Literature on the various schools of interpretation is growing rapidly, as is the debate on their relative merits. The most recent work in this country is that of Gerald Press. See both *Plato's Dialogues: New Studies and Interpretations* (Lanham, Md.: Rowman & Littlefield, 1993) and *Who Speaks for Plato? Studies in Platonic Anonymity* (Lanham, Md.: Rowman & Littlefield, 1999). Francisco J. Gonzalez, ed. *The Third Way: New Directions in Platonic Studies* (Lanham, Md.: Rowman & Littlefield, 1995) is another new collection of essays attempting to forge a clear and fruitful interpretive path integrating two schools of Platonic scholarship that often seem to be at loggerheads. These scholars are responding to the chasm between the Anglo-American analytic school prominent and influential in Oxford since the 1960s, as well in the United States, and other schools generally originating in Europe—notably, the Tübingen school, which is responsible for a body of literature supporting the claim that Plato's unwritten doctrines must be extracted from his written works, and that these doctrines are at the core of his philosophy. A second group of scholars radically opposed to the analytic approach usually ignores arguments in the dialogues as tangential; these scholars approach the dialogues from the critical standpoint of literary theory, studying the works as dramas or, in more recent terminology, enactments. Giovanni Reale and Thomas Schlezak are important examples of scholars who espouse individual versions of the latter sort of view.

48. See Julia Annas, *Platonic Ethics, Old and New* (Ithaca, N.Y.: Cornell University Press, 1999), 4.

49. L. Campbell, *The Sophistes and Politicus of Plato* (Oxford: Oxford University Press, 1867).

50. They rely upon certain presumptions: first, that Plato (as any other author) has a prose style with some features of which he is conscious and thus can control, and others of which he is unconscious and therefore is not able to control; second, that we can determine which features are unconscious; and third, that these are the only proper subjects of stylometric studies. See the excellent and concise discussion in T. M. Robinson "Plato and the Computer," *Ancient Philosophy* 12 (1992): 375-82.

51. See, for example, DL III 37 on the *Laws* as the final work of Plato.

52. Gregory Vlastos provides the best-known contemporary examples of this strategy.

53. Among the difficulties are those of sorting out different styles of prose as well as sections of poetry and even quotations: should the word counts include all of these or not? If, for example, the speech of Lysias at *Phaedrus* 230e6-234c5 is included in the count, two problems arise: Is the speech indeed a quotation of Lysias or is it Plato's masterly stylistic imitation of him? And because the speech is consciously shaped, perhaps linguistic usage departs from Plato's usual style, rendering the passage inappropriate for inclusion in stylometric analysis.

54. See note 49 for a listing of recent publications and reviews with extensive arguments concerning the matter of chronology in Plato.

55. Plato: *Complete Works*, (Indianapolis, Ind.: Hackett, 1997), xiv.

56. Perhaps the most thorough recent study was published in 1995: Debra Nails' *Agora, Academy, and the Conduct of Philosophy*, (Dordrecht: Kluwer). Nails devotes chapters 4-7 to a thorough examination of the stylometric analyses of Ledger and Thesloff, eventually concluding that no stylometric methods will untangle Plato's chronology. Taken together, the inconclusiveness of statistics on hiatus, clausulae, and certain key particles and other expressions that are thought to be unconscious and thus accurate indicators of chronological period, along with the possibility that Plato edited his work repeatedly, as Dionysius of Halicarnassus remarks in *De compositione verborum* 25. 210-12, denies us an absolutely sure chronology, according to Nails. Recent stylometric studies include Gerard R. Ledger, *Re-counting Plato: A Computer Analysis of Plato's Style* (Oxford: Oxford University Press, 1989) and Leonard Brandwood, *The Chronology of Plato's Dialogues* (Cambridge: Cambridge University Press, 1990). Brandwood also summarized his findings two years later in *Stylometry and Chronology, Cambridge Companion to Plato* (Cambridge: Cambridge University Press, 1992): 90-120. Reviews by P. Keyser of Ledger (*Bryn Mawr Classical Review* 2 [1991]: 422-27) and of Brandwood (*BMCR* 3 [1992]: 58-73) conclude that Brandwood is successful in providing a chronology of early, middle, and late groups without chronology within the groups. Ledger, however, does not unequivocally confirm Brandwood's findings. See also Debra Nails, "Platonic Chronology Reconsidered" (*BMCR* 3 [1992]: 314-27) and T. M. Robinson, "Plato and the Computer," *Ancient Philosophy* 12 (1992): 375-82. Another detailed discussion of Ledger and Brandwood is that of Charles M. Young, "Plato and Computer Dating," *Oxford Studies in Ancient Philosophy* 12 (1994): 227-50.

57. See Vlastos, *Socrates, Ironist and Moral Philosopher* (Ithaca, N.Y.: Cornell University Press, 1991), 46-47. The grouping of the dialogues he accepts is as follows: Group I includes the Elenctic Dialogues (in alphabetical order) the *Apology*, *Charmides*, *Crito*, *Euthyphro*, *Gorgias*, *Hippias Minor*, *Ion*, *Laches*, *Protagoras*, *Republic* I, and Transitional Dialogues *Euthydemus*, *Hippias Major*, *Lysis*, *Menexenus*, *Meno*; Group II includes (in probable order) *Cratylus*, *Phaedo*, *Symposium*, *Republic* II-X, *Phaedrus*, *Parmenides*, *Theaetetus*; Group III includes *Timaeus*, *Critias*, *Sophist*, *Politicus*, *Philebus*, and *Laws*.

58. See Irwin, *Plato's Ethics* chapter 1, Fine, Kraut, "Introduction" in *Cambridge Companion to Plato*. White (*Plato on Knowledge and Reality* [Indianapolis, Ind.: Hackett, 1976], 24 n. 22, 27 n. 50, 117), and Guthrie (*History of Greek Philosophy* Vol. IV, 39-56) also suggest certain internal orderings for specific groups of dialogues. Charles Kahn's recent extended argument for taking at least the Socratic dialogues as a coherent corpus (*Plato and the Socratic Dialogue: the Philosophical Use of a Literary Form* [Cambridge: Cambridge University Press, 1996]) attempts to synthesize a chronology with Plato's literary plan. For criticism of Kahn's approach, see Charles L. Griswold, Jr., "E Pluribus Unum," *Philosophical Review* 105 (1996): 1-32.

*ribus Unum?* On the Platonic ‘Corpus,’” *Ancient Philosophy* 19 (1999): 361-97, as well as Kahn’s “Response to Griswold,” *Ancient Philosophy* 20 (2000): 189-93. Griswold has the final word in “*E Pluribus Unum?* On the Platonic ‘Corpus’: the Discussion Continued,” *Ancient Philosophy* 20 (2000): 195-97.

59. See Christopher Shields, “Simple Souls” (this volume), and Kahn, *Plato and the Socratic Dialogue*.

60. This position does not deny, however, the innate plausibility of such claims as Hackforth’s contention that the metaphor of soul as charioteer and horses found in the *Phaedrus* is likely to have followed rather than preceded the extensive argument for tripartition of *Republic* IV. It is more reasonable that Plato would have introduced the metaphor to an audience already acquainted with the model of soul in three parts, since he does not motivate in the *Phaedrus* the distinctive tractability of the white horse, the unruliness of the black horse, or the unswerving reason of the charioteer.

## I. The Nature of the Soul

## Fallacies in the *Phaedo* Again

Cass Weller

### Introduction

Treatments of the final argument for the immortality and imperishability of the soul in the *Phaedo* (102a-107b10) generally divide into two sorts. According to one, (A), an individual soul is an immanent character of the form Soul. It is a unit property or trope which mediates between the universal and the concrete individual of common sense. On the other view, (B), the soul is, to put it anachronistically, a Cartesian substance. An individual soul is the subject of psychological attributes such as thought, perception, and desire and coincides with what is ordinarily thought to be a person or Cartesian ego.

Gallop in his commentary accordingly provides two schematic reconstructions of Plato's argument.<sup>1</sup> He clearly favors the Cartesian view, (B), and for reasons with which I generally concur. Since he is dealing at the level of families of interpretation his commentary does not provide us with a detailed criticism of any of the members of the other family of interpretations (A). I propose to do just that, not simply in order indirectly to defend the Cartesian interpretation, but in the belief that deeper understanding of Plato's argument will result from such an exercise. In particular I hope to show that the mistake in Plato's reasoning is not the one imputed to him by either (A) or (B).

On the immanent character interpretation (A), it is charged that an individual soul is not of the right logical type to be (predicatively) alive and immortal and so can't be inferred to be imperishable from the putative entailment from immortality to imperishability. On the

Cartesian view, (B), Plato is typically seen to be guilty of the generic fallacy common to ontological arguments whose premises are conceptual truths and whose conclusions are or entail existential statements. Accordingly, the argument moves illicitly from the premise that the essence of soul includes the property of immortality, which in turn includes imperishability, to the conclusion that there are souls and they are imperishable. However, these premises only entail that if  $x$  is a soul then it is imperishable. It would take additional argument to establish that there are souls which in fact are always alive and therefore imperishable.

The specimen of immanent character interpretations I will examine is the one defended by David Keyt (1963).<sup>2</sup> According to this interpretation Plato commits two serious fallacies in the course of the argument (*Phaedo* 102a-107b) purporting to establish that the soul is immortal and imperishable. (1) He is guilty of equivocation because he shifts from the technically contrived sense of  $\alpha\vartheta\alpha\varpi\tau\omega\varsigma$ , “undead” (“alive”) back to its ordinary sense, “immortal.” (2) He commits the fallacy of composition by illegitimately inferring from the premise that what is immortal (has life everlasting) is imperishable to the conclusion that the soul is imperishable. This is fallacious, since at most what he is entitled to infer, even winking the equivocation, is that the ensouled body—not its soul—is imperishable. Keyt’s charge is that the soul doesn’t animate a body by being an instance of the property of being alive or the form The Alive Itself. Instead, the soul animates a body by being present in it in the manner in which an immanent form of The F is present in the many Fs. That is, on Keyt’s formulation (169) Plato treats the soul in Socrates *as if it were* an immanent form logically subordinate to the immanent form of life. Thus the statement “What has a soul in it is alive” does not entail “The soul is alive”; and to suppose that it does is to commit the fallacy of composition in one of its forms. The one he has in mind is confusing the relation of class inclusion with class membership and confusing the relation of logical subordination as an entailment relation with predication. Thus, according to Keyt, for Plato to infer from the logical relation of subordination between immortality and imperishability that the soul is (predicatively) imperishable would be tantamount to inferring that the concept of a horse is (predicatively) an animal because the concept of a horse is logically subordinated to the concept of a mammal and the concept of a mammal is logically subordinated to the concept of an animal.

With appropriate qualifications I concur with (1), the charge of equivocation on  $\alpha\vartheta\alpha\varpi\tau\omega\varsigma$ . However, I dissent from (2). The reason

is that I think it a mistake to construe the soul’s being alive along the lines of the tallness of the-tall-in-Socrates. Rather, I think on balance that it would be better to construe the sense in which the soul is alive and can’t admit of death as parallel to the sense in which ordinary snow is cold and can’t admit of being hot. On this interpretation, the Cartesian one, the soul brings life to the body it inhabits by, as it were, infecting the body with the very property it instantiates. It is this, I believe, Plato intends. And though his argument for it can be criticized as committing a fallacy of some kind, it is not the one Keyt charges him with.

One of the strengths of Keyt’s interpretation is that the idea of items that flee or perish at the approach of something involving relevant contrariety is introduced in the context of the immanent characters, the tall-in-us and the-short-in-us (102d5 ff.). And so it is natural to assume that in the sequel wherever there is mention of some item being compelled to flee or perish at the approach of contrariety in some mode or other we are dealing with characterlike items rather than with the owners of such characters. Thus, even snow and fire are to be regarded as characters or metaphysical constituents of ordinary fire and snow and not as the ordinary causal agents of everyday life (Keyt 168 and n. 2); and this despite the fact that one would have thought it much more natural to suppose that it’s the ordinary snow in my hand that perishes and the flickering flames that die. This sweeping assumption or at least its application in the case of the soul receives some support from the following considerations. A soul is in a body but not as wine is in a glass; so it might seem natural to suppose that the relevant sense of “in” in the case of the soul is rather like that of “in” in the case of immanent characters. The mathematical examples would seem to offer corroboration. Consider, for example, that the crucial question of what makes a body alive is introduced on the heels of the prompt (105c4-6) that the presence of the unit (monad), which seems more propertylike than propertied, makes what it is present in odd. Moreover, earlier at 104d5-7, where Plato is presumably speaking more precisely, we are told, in the same language of the crucial answer, that whatever the Form ( $\iota\delta\epsilon\alpha$ ) of three takes possession of ( $\kappa\alpha\tau\alpha\sigma\chi\eta$ ) is compelled to be not only three but odd.

### Preliminary Objections and Observations

As against all this I begin with a modest observation. If souls are virtually indistinguishable from immanent characters or tokens of THE-

ALIVE-ITSELF then the distinction between the clever  $\alpha\tau\tau\iota\sigma$  and simple  $\alpha\tau\tau\iota\sigma$  collapses where the explanandum is what makes a body alive. The distinction depends on there being two kinds of answer to the question, “What makes what is F, F?” where “F” is a predicate standing for a member of a pair of opposites. The naive ( $\epsilon\bar{\nu}\eta\theta\omega\varsigma$ ; 100d4,  $\alpha\mu\alpha\theta\bar{\eta}$  105c1 unlearned) and safest ( $\alpha\sigma\phi\alpha\lambda\epsilon\sigma\tau\alpha\tau\omega\varsigma$ ; 100d8, 105b7) answer is that F-ness makes what is F, F; whereas the more sophisticated answer has it that the presence of something  $x$  which, while not itself F-ness, is essentially characterized by F-ness is logically sufficient for something  $y$  being F. Thus, the presence of The Hot explains in the safest manner why a body is hot, while the fact that it is on fire explains in more sophisticated terms why it is hot, i.e., a nonlogical causal factor. Therefore, since the presence of a soul in a body is explicitly introduced as the more sophisticated explanation of why a living body is alive (105c9-11), a soul has to be distinct from the transcendent form of The Alive or any of its immanent characters. Otherwise, the sophisticated and illuminating answer gets dimmed down to “What is alive (a living body) is alive by THE ALIVE or by what is logically subordinated to THE ALIVE.”

Another modest observation. If the soul is more like the property of being alive than like what has that property then how could it be an object of concern and care? Of course, one might accept this observation and just add a charge of *ignoratio elenchi* to the list.

We must of course recognize that souls are two-faced. They are unwilling or self-deceived denizens of the sensible realm of becoming, even as they aspire to the intelligible realm of being. It is not at all surprising that souls bear certain similarities to Forms, transcendent or immanent. Plato has already been at pains (78c-80c; especially 80b1-5) to make a case for such similarities. But that very fact shows that Plato was also well aware of differences between souls and forms. In particular it suggests that Plato recognized that the relation between a soul and the body it occupies is *sui generis* and not to be subsumed under the relation of form exemplification. So we needn’t suppose that he illicitly traded on the similarities, if there’s a better explanation for the illicit inference. Once again, however, one might appropriate the observation that souls are likened to forms and take that to explain the attractiveness of the bad inference which he is consequently unable to resist.

One final opening observation. Keyt assumes—and it is crucial that he do so<sup>3</sup>—that in the *Phaedo* Plato posits forms corresponding to substantives in addition to forms corresponding to adjectival opposites. Not only are there such forms as THE HOT, THE BEAUTI-

FUL, THE JUST, THE ODD, THE ALIVE, and so on, but also such forms as THE SOUL ITSELF, SNOW ITSELF, and presumably forms corresponding to every natural kind term.<sup>4</sup> However, the sample lists of forms scattered throughout the *Phaedo* include only forms of opposites.

The interpretative stance I take is minimal. I willfully ignore relations that the Platonic theories of the *Phaedo* bear to Platonic theories in other dialogues. In particular I will not seek to illuminate the arguments of the *Phaedo* by appealing to the metaphysics of the *Timaeus*. It seems to me that what the argument for the soul’s immortality is, and what goes wrong with it, can be understood without engaging in the debates surrounding Plato’s philosophical development.

### The Heart of the Matter

I now want to focus on the claim central to Keyt’s interpretation, namely, that the snow that can’t stand the heat is the immanent character or form of snow. By undermining it I hope to slow, if not derail, the train that leads to a similar treatment of the soul. The obvious place to break into Plato’s chain of reasoning is at 102d5. Here we are introduced to the so-called immanent characters. Not only will The Large Itself not suffer itself to be small; the-large-in-us won’t accept the small. Whenever the small approaches it, it retreats or perishes. There then (103a4-b9) follows what seems to be a minor digression. One of the auditors gets confused and thinks that the doctrine under discussion contradicts the earlier discussion in which Socrates argued that opposites come to be from opposites. They had agreed that the hot becomes cold, that the small becomes large. And now we hear that the-small-in-us will never become large and by extension the-hot-in-us will never become cold. Socrates then carefully distinguishes between, on the one hand, the forms of opposites, both transcendent and immanent, and those objects that have them and are named after them, on the other hand. For example, the large is to be disambiguated into (1) that which is (predicatively) large (Mt. Rainier) and (2) that which makes what is large large. The latter is further divided into (2a) The Large that mediately makes Mt. Rainier large, i.e., the transcendent form THE LARGE and (2b) the Large that directly makes Mt. Rainier large, a corresponding immanent character.

With a clear distinction between what is F and The F-that-makes-it-so (1 and 2 above) in hand, Socrates again takes up the thread of the argument. He proceeds to distinguish between fire and the hot,

snow and the cold, fully expecting that the auditor will hear this as the distinction between fire as what is hot and THE HOT as what makes it so and between snow as what is cold and THE COLD as what makes it so. We are then told (103d5-8), with reference to the earlier flee-or-perish passage, that never will snow which has received the hot attack be the very thing that it is, namely, snow, as well as hot; but when the hot attacks, it (the snow) will retreat or perish.

Now so far I see no reason to let the reference back to the earlier flee-or-perish passage turn snow into the immanent form of snow. We are still under the rule of the distinction between what is F, which in this case happens to be snow, and that-which-makes-it-F, i.e., The F, immanent or transcendent.

At 103e2-5 he generalizes:

It stands with such things that not only does the form itself (The F) deserve its name forever, but further that something else, x, which is not the form, nevertheless, has its character ( $\mu\sigma\varphi\eta\gamma$ ) whenever it, x, exists. [my translation]

Now this looks for all the world to be a simple generalization of the claim that The Cold deserves its name forever (The Cold IS cold) and snow, which is not The Cold, nevertheless, has the character ( $\mu\sigma\varphi\eta\gamma$ ) of The Cold, whenever it exists. Therefore, since the soul is also argued, via a causal likeness inference, to be essentially characterized by the form of an opposite, it too is one of those things which is not the form of an opposite, i.e., THE ALIVE ITSELF, nor the corresponding immanent character, but rather something which has that character whenever it exists. The case for there being an immanent form of snow has not been made, and thus the case for there being an immanent form of soul finds no support here.

But let's suppose, anyway, in keeping with Keyt's view, that in our passage what has the character of The Cold, whenever it exists, is not ordinary snow, but rather the immanent counterpart of the transcendent Snow Itself. The supposition implies a disjunction: either (i) the immanent character-snow is something which is cold and made so by THE COLD ITSELF or (ii) the immanent character-snow is always accompanied by the immanent character-cold.<sup>5</sup>

Disjunct (i) is much more likely given the stress on having rather than being the character-cold. Accordingly one would be faced with the puzzling idea of one immanent character being present in another. This would raise the further embarrassing question of whether this snowball was cold because of the presence at a second remove of the immanent character-cold in it, i.e., by the presence of the immanent

character-cold in the immanent character-snow itself present in the snowball, or because of a first level presence of another, numerically distinct, immanent character-cold in the snowball. I thus take this to be a dead end.

On disjunct (ii), the less likely, the immanent character-snow is only accompanied by the immanent character-cold, so that these two characters are present in the snowball at the same level. Presumably then, at the approach of the hot whatever befalls the immanent character-cold befalls the immanent character-snow; they both either flee or perish. But so too the snowball. Thus if the character-snow is essentially accompanied by the character-cold, the snowball essentially has the character-cold. What is crucial to notice here is that there is no fallacy of composition in inferring that snowballs are essentially cold. So if we accept what, in any case, lacks textual warrant, that the snow in the passage in question is the immanent character-snow, we get either the unpalatable consequences of (i) or the benign consequences of (ii) which fail to yield a fallacy of composition.

The crux lies in how this applies in the case of the soul. One would naturally suppose that Keyt understands the immanent character-snow to be the analogue of the soul. They are both equally subject to the flee-or-perish principle on his understanding of that principle. Moreover, it yields the results he argues for; namely, that in the one case Plato confusedly treats as a relation of predication either (a) the relation of accompaniment between the elements in a complex immanent character, (b) the part-whole relation between a complex immanent character and one of its elements, or (c) the relation of logical subordination between two attributes.

I suspect, however, that Keyt does not fully subscribe to the analogy as I would understand it. I have been assuming that corresponding to tokens of the immanent character-snow there are snowballs each of which is a ball of snow because of the presence in it of an immanent character-snow; and thus, somewhat disingenuously, that corresponding to souls as immanent characters are souls which have them. Keyt might well dissent on two counts. (1) He may deny that there is a distinction of logical type between this snowball and an immanent character-snow on the ground that in the *Phaedo* ordinary spatio-temporal continuants just are leaky bundles of immanent characters in the  $\delta\pi\delta\delta\chi\eta\gamma$ . (2) He may, without invoking the doctrine of congeries, simply claim that the analogy is defective in that souls as immanent characters, unlike immanent characters of snow, have no eponymous counterparts.

All this notwithstanding, let's push ahead and complete the exercise and see if by assuming immanent forms of soul for souls in strict

parallel with immanent forms of snow for portions of snow we reproduce the previous disjunctive result of either (i) the impasse of extremely puzzling metaphysical questions or (ii) the absence of the fallacy of composition.

Suppose then that the immanent form of soul has the immanent character of the alive. We have the parallel difficulty of one immanent character being present in another. We also have the parallel question of whether the soul is alive by having present in it, at different levels, two numerically distinct tokens of the-alive; and further whether we have need of yet a third token of the-alive to account for a living body; or whether one can do the whole job. That is, suppose that the soul is alive by having an immanent form of soul which in turn has a token of the-alive and that a given animal is alive because it has a soul which has an immanent form of soul which has a token of the-alive.

To complete the exercise let us now suppose, on the other hand, in accordance with reading (ii) above of 103e2-5, that an immanent form of soul is always accompanied by an immanent character of the-alive. Here as in the parallel case, *mutatis mutandis*, it follows, without fallacy, that the soul is always alive whenever it exists, on the general assumption that if an immanent form of F is always accompanied by an immanent form of G, then what is predicatively F is predicatively G.

As I indicated earlier, the foregoing treatment is not altogether in the spirit of Keyt's interpretation, which after all depends on virtually identifying the soul with the immanent character-the-alive and denying that there are eponymous counterpart Cartesian souls. He will opt for reading (ii) despite the fact that what flees or perishes has the character of an opposite. But notice, if this is the way the wind blows and the putative statement "Souls are alive," if taken predicatively, is a category mistake, then so too the putative statement "Snow is cold."

It must, however, be conceded that the assumption I have been operating under is not without liabilities. If there are ordinary eponymous portions of snow and corresponding eponymous souls I am committed to the unnatural sounding idea that individual souls are in some sense analogous to individual snowballs. But snowballs don't take possession of bodies in any way remotely analogous to souls' taking possession of bodies. Nor do we rely on a causal likeness inference to establish that snow is cold.

It is no doubt some such considerations that would lead Keyt to suppose that it is the ensouled body which is the proper analogue to the snowball. Let us then suppose this to be the controlling analogy of

Keyt's interpretation,<sup>6</sup> notwithstanding its lack of direct textual support. If the ensouled body is analogous to the snowball, then the soul as immanent form of living animal will be the analogue of the immanent form of snow. Notice that this analogy also suggests that snow and ensouled bodies are composites of their respective concrete materials and essential immanent characters. Here's the picture we get.

$$\begin{array}{ll} \text{SNOW} = & \text{WATER} + \text{an immanent character of the cold} \\ \text{LIVING ANIMAL} = & \text{BODY} + \text{soul} \end{array}$$

Putting things this way shows why in one respect his interpretation is attractive. When a snowball perishes by heat there is a corpse, the puddle of water, just as when an animal perishes. In each case what is no longer present is a critical immanent character, whether or not the character itself is destroyed. What else is there for the soul to do except to play the role analogous to the immanent character of the-cold? But notice the further commitments of this analogy. It leads to treating a living animal as a phase of a body, which further leads one to suppose that corpses can be regularly revived; after all the puddle can be refrozen, even if not by the same token of the-cold. Snow as a composite is essentially cold (frozen); the water in that composition is cold (frozen) too, though not essentially.

However, let us provisionally accept the analogy this far.

This ball of snow, which is frozen water, is such that it needn't be snow [because the water which is frozen needn't be].

This animal, which is an ensouled body, is such that it needn't be an animal [because the body which is ensouled needn't be].

What is crucial here is that just as the predicate "frozen" or "cold" is contingently true of portions of water, the predicate "ensouled" is contingently true of bodies. One would therefore conclude that the predicate "alive" is also contingently true of bodies. Even if we regard "x is ensouled" as a metalevel statement to the effect that an immanent character of the-alive is in x, it still implies that x is (predicatively) alive. For I am assuming that for any immanent character c named after The F Itself or always accompanied by such an immanent character, if c is in an individual x then x is predicatively F. Recall, however, that the mistake Keyt finds in Plato's argument is inferring that the soul is alive from the fact that the composite in which the soul is a vivifying element is alive, whereas only the composite is the proper subject of the attribute being alive. Of course, if only the com-

posite of body and soul is the proper subject of the attribute being alive, then the controlling analogy of Keyt's interpretation, that between snow and the composite of body and soul, breaks down. For as we have just seen being cold is true of both the composite snow as well as one of its elements, the water that happens to be crystalline.<sup>7</sup>

Let me conclude this section by bringing to the fore a different analogy that plays more readily into my hands and that's closer to Plato's introduction of the soul as the more sophisticated explanans of a living body. Fire, we are told at 105a1 and b8-c2, is one of those things that is always accompanied by an opposite, in this case THE HOT, and subsequently brings it to whatever it occupies making what it occupies predicatively hot. Now whether or not one supposes that Plato is further committed to the existence of a Form of fire or to corresponding immanent characters, one can understand perfectly well at the level of common sense the idea of fire as a causal agent which makes things hot by contagion. Plato may have independently held that the presence of fire in any ordinary thing is also necessary for that thing being hot. But in any case it's fire as what everyone is familiar with that Plato is supposing is hot and sufficient for making other things hot by being present in them or taking possession of them in some way or other.<sup>8</sup> Fire as elusive, flickering, and not quite a body among others yet capable of heating them seems just the right analogue for the soul. As an element present in bodies, especially ones not in flames, fire is at least as seductive a model as immanent characters for understanding how souls are present in bodies. Moreover, it's easy to see how one might conclude from the fact that fire, as what heats, is itself hot that the soul, as what vivifies, is itself alive.<sup>9</sup>

### Recapitulation

Let's begin again and grant for the time being that Plato has established that the soul can't admit of death just as snow can't admit of heat and fire can't admit of cold. Now to conclude from this that the soul is immortal or deathless is legitimate only to the extent that "immortal" or "deathless" means "can't be dead" as understood in this context. A soul is immortal only in the sense that there is no such thing as a dead soul, just as there is no such thing as a cold fire or hot snow or an even triplet. It is essential to each soul that it be alive.

If, however, Plato is using "immortal" or "deathless" in its ordinary sense of "has life everlasting," he is not entitled to conclude that the soul is immortal. From the previous claim that

*It is essential to each soul that it be alive.*

it follows that

*If an object x is a soul at a given moment of time, it is alive at that moment.*

It also follows (trivially) that

*If an object x is a soul at every moment of time, then it is alive at every moment.*

However, it does not further follow that

*If an object x is a soul at a given moment t then it is alive not only at t but at every moment subsequent to t.*

But this, at the very least, is what is implied by the immortality of the soul conventionally understood as having life everlasting.

What is tricky in the argument is not simply a gross equivocation on the two senses of "immortal" or "deathless." It is rather more complicated. Ordinarily, to say that something which was once alive is dead is to say that it no longer exists. That (pointing to the lifeless body) used to be my pet goldfish, Bill Casey, but it died of a mysterious brain tumor. It goes without saying that Bill the goldfish no longer exists, even though the corpse, which hasn't yet decomposed, does still exist. Conversely, to say that Bill Casey the goldfish, which was once alive, no longer exists is to say that it is dead. If something that was alive no longer exists, then it is dead. But it follows that if something that was alive is not dead, then it still exists. So if the goldfish which was alive yesterday is not dead, it still exists.

Now suppose we say that the goldfish will never admit of death and remain what it is, namely, a goldfish. A dead goldfish isn't a goldfish. It no longer exists. In a certain quirky sense, then, the living goldfish will never be a dead goldfish. It may become or turn into a dead goldfish by ceasing to exist, but the goldfish itself will never be a dead goldfish. At best, a dead goldfish is a former goldfish. We have thus conferred a kind of immortality on the goldfish.

Consider the case of the soul. It will never admit of death and remain what it is. A dead soul is not a soul. The problem in this case, however, is that it's hard to make sense of a soul, which as a soul must be alive, turning into a dead one; there is no analogue of the corpse. So if a soul were to lose its life and thereby cease to exist there would

be no transformation of one thing into another but rather an absolute going out of existence. But that's absurd. Nothing which in the ordinary course of its existence undergoes internal changes, as a soul does, can just go out of existence altogether without a trace, without being transformed into something else. So not only can't a soul be dead, it can't become a dead one. So it will be alive for ever.

If the preceding argument strikes one's ear as unabashedly Aristotelian, I make no apologies but claim instead that what one is hearing is Aristotle's Platonism, absent the hylomorphism.

In any case, a reasonable objection to this line of reasoning is to reject the conclusion by denying the premise that the soul is (predicatively) alive; that is, deny that the concept of being alive applies unequivocally to the soul. The point is this. For the concept of being alive to be true of an object of some kind K it must be possible for there to be a dead K, which is to say, individuals of kind K must be composite. It is not possible for there to be a dead soul. Therefore, no soul is, strictly speaking, alive.

It is something like this point, I believe, that Keyt is responding to when he charges Plato with the fallacy of composition. For Keyt allows that at most the entailment from immortality to imperishability establishes that the ensouled body is imperishable. Thus he supposes that only the composite of body and soul is the proper subject of the attribute being alive. And for Plato to conclude that one of the elements of this composite, even the one in virtue of which the composite is predicatively alive, is itself predicatively alive, is to commit the fallacy of division or composition depending on one's point of view. Notice now how Aristotelian Keyt's position is. According to his Plato the soul is the form of a body which is essentially alive and only the composite of body and soul is predicatively alive. For Plato to mistakenly infer that souls are predicatively alive is for him to misunderstand this Aristotelian strand of his own thought.

On both of our views, then, Plato relies on what I have been calling a causal likeness inference, "Because of  $x$ ,  $F$ s are  $F$ ; so  $x$  is  $F$ ." However, there is a telling difference. For Keyt the inference occurs at the metaphysical level, so to speak: "because of some immanent character  $x$ , living things are alive; so  $x$  is alive." Here  $x$  is of the wrong logical type to be predicatively alive. It does not instantiate properties at the same level at which living things and ordinary objects do. However, on the view I have been urging, the inference occurs at the lower level; "because of souls, which are first level individuals, living bodies are alive; therefore souls are alive." What I claim is that Plato's reasoning would be good, if the conclusion he drew had the sense of "souls are eminently alive," made familiar to us by Des-

cartes.<sup>10</sup> That is, a soul makes a living body alive due to its causal powers grounded in its intrinsic features. A soul is predicatively natured in whatever way it has to be in order to have the causal powers it has. But Plato misconstrues this and thinks that for a first level individual to be eminently  $F$ , that is, to be predicatively what it takes to make things  $F$ , is for it to be predicatively  $F$  in a superior way. This is quite different from the sort of level confusion which Keyt charges and which leads to problems of self-predication. Of course, the mistake I have discerned would also recur all the way up, if in fact Plato holds that the form  $F$ -ness is predicatively  $F$  because it makes an  $F$  predicatively  $F$ . In this case, however, the mistake would be further compounded, since from an Aristotelian standpoint one could come to infer mistakenly that a form that is eminently  $F$  is also predicatively  $F$  only by having first mistakenly supposed that it is eminently anything. However, whether Plato believes that The  $F$  Itself is predicatively  $F$  and just how to construe the causal likeness inference in the case of forms are the enduring questions of Plato's metaphysics.\*

## Notes

1. *Plato*, *Phaedo*, translated with notes by D. Gallop (Oxford: Clarendon, 1975), 192-222.
2. David Keyt, "The Fallacies in *Phaedo* 102a-107b," *Phronesis* 8 no. 2 (1963): 167-72.
3. He can avoid this assumption only by supposing instead a reductive analysis of ordinary objects in terms of congeries of immanent qualitative characters. For example, a living animal would be an unstable bundle of tokens of qualitative and quantitative forms including a token of THE-ALIVE-ITSELF. But such a supposition is even less warranted by the text than the assumption of forms for natural kinds, especially when one considers that Plato bothers to examine in some detail other exotic reductive claims about the soul, for example, the attunement doctrine.
4. The assumption has not gone unchallenged (See R. G. Turnbull, "Aristotle's Debt to the Natural Philosophy of the *Phaedo*," *Philosophical Quarterly* 8 [1958]: 131-46, and A. Nehamas, "Predication and Forms of Opposites in the *Phaedo*," *Review of Metaphysics* 26 [1973]: 461-94).
5. Notice the puzzling question this case gives rise to. If there is a form of SNOW ITSELF and it is definable in terms of the forms WATER ITSELF and THE COLD ITSELF, then it is reasonable to suppose that the immanent character of snow is a composite of immanent characters of water and the cold. In that case, however, what's the point of saying that the immanent character-snow, one

of whose constituents is an immanent character the cold, is also accompanied by (another?) immanent character of the cold?

6. Notice that a previously observed perplexity (cf. n. 5 above) has still not been dissipated. If there is an immanent character of snow then there is a transcendent form Snow Itself. The latter is presumably to be defined in terms of, *inter alia*, the form Water Itself and the form The Cold Itself. This implies that the immanent characters corresponding to Snow Itself are packages of the immanent characters corresponding to Water itself and The Cold Itself. In that case an immanent character of The Cold Itself, which is either predicatively present in or simply accompanies an immanent character of Snow Itself, would appear to be an extra immanent character with no work to do. At best to speak of the predicative presence or mere accompanying presence of the immanent character of The Cold Itself would be a highly misleading way of describing an essential element of an immanent character of Snow Itself.

7. One could reply that “alive” is after all true of the body as an element of the composite but only contingently; whereas it is true of the composite essentially. But this simple modal difference doesn’t go far enough, if we take seriously the idea that an animal is a composite of body and soul. For if a part and the whole of which it is a part are distinct entities, then on the Platonic theory under consideration each will have its own numerically distinct immanent character of the-F, if each has the same attribute and is predicatively F. Thus the animal will be alive, albeit essentially, in virtue of one soul and immanent character of the-alive and its body will be alive in virtue of another soul and immanent character of the-alive. Many of these difficulties are avoided, if one regards the soul as an individual of a different logical type than that of an immanent character and treats the composite animal as a unity of two quasi-concrete individuals of the same logical type, a unity to which alone the predicate “alive” applies.

8. The pattern of inference governing the passage is expressed in the following pattern of question and answer:

Any  $x$  in which *what* is present is such that  $x$  will be F?

Any  $x$  in which  $K$  is present is such that  $x$  will be F.

Where “F” is replaced with “hot,” “K” is replaced with “fire.” Where “F” is replaced with “alive,” “K” is replaced with “(a)-soul.” The fact that any body in which a soul is present is such that it is a living body is one from which Plato implicitly infers, on the strength of 103e2-5, that the soul is essentially characterized by the form of the opposite which it imposes on the bodies it occupies. We might state the principle drawn from these two passages in one of two ways.

There are certain kinds, K, such that every member  $x$  of K, while not a form of an opposite, F-ness, is always characterized by F-ness as long as it exists and if  $x$  is present in any  $y$ ,  $y$  will also share in F-ness because  $x$  does so essentially.

If any  $y$  is characterized by the opposite, F-ness, because of the (non-predicative) presence in it of  $x$ , then  $x$  is a member of kind K each of whose members is characterized by F-ness for as long as it exists.

Each version constitutes an extension of the safe and simple causal likeness principle: If  $x$  is F then there is something, F-ness, because of which  $x$  is F due to the fact that F-ness is F.

9. It seems to me not an obviously foolish question whether fire is hot in the same sense as what it heats. Of course, Plato, I should think, would say that there are different senses or degrees and claim that fire is more truly hot than what it heats.

10. There is a closer source. I think the idea is already implicit in the Aristotelian thesis that a doctor’s ability to make patients healthy entails that the doctor is healthy in virtue of his knowledge of health.

I am much indebted to David Keyt for generous comments on several drafts of this paper, above all for his patience in revisiting a paper he wrote more than thirty years ago. I also wish to thank Marc Cohen and Jean Roberts for helpful suggestions.

## The Arguments in the *Phaedo* Concerning the Thesis That the Soul Is a *Harmonia*

C. C. W. Taylor

At *Phaedo* 85e-86d Simmias puts forward an argument to show that the soul cannot be immortal. The premises of the argument are firstly that the soul is a *harmonia* of the elements that compose the body, the hot, cold, wet, dry and so on (86b6-c2), and secondly that no *harmonia* can exist unless the elements of which it is a *harmonia* maintain the proper interrelation.<sup>1</sup> (This point is made in 85e3-86b5 with reference only to a particular case, the *harmonia* of a lyre, but is clearly to be taken generally.) It follows that when the interrelation of the bodily elements has been dissolved by death, the soul-*harmonia* cannot exist apart. This argument is presented in the dialogue as posing a major objection to the thesis of the immortality of the soul; those who had been convinced by Socrates' previous arguments are now thoroughly dismayed (88c). It is, therefore, worth some consideration, particularly since the premise that the soul is a *harmonia* expresses a philosophical doctrine whose sense is far from clear. Furthermore, the counterarguments by which Socrates claims to refute Simmias have provoked considerable disagreement among commentators as to their interpretation, while questions may be raised as to their validity. I propose, then, first to ask what is the meaning of the thesis that the soul is a *harmonia* and second to examine Socrates' arguments against Simmias.

The ambiguity of the thesis that the soul is a *harmonia* emerges from

consideration of the different shades of meaning that the word *harmonia* may have. Formed from the verb *harmozein* ("fit together") it expresses the idea of things being fitted together in an exact arrangement to make an integrated whole, but particular uses express various aspects of the basic sense. Thus the word sometimes means "proportion," particularly in contexts where elements are mingled in proportion, as where Empedocles describes painters mixing their colors "mixing them in proportion, some more and some less" (DK 31 B 23, line 4), and sometimes "arrangement" or "organization" (conveying the idea of the proper relation of parts), as when Heraclitus refers to the *harmonia* of opposite forces in a bow or a lyre (DK 22 B 51). Or again a *harmonia* may be identical with a complex of parts in a certain order or arrangement; this is the sense in which the word can mean "joint" or "framework" (see Liddell and Scott). Aristotle's discussion of the soul-*harmonia* thesis at *De Anima* 1.4 deals certainly with the first two senses, and perhaps also with the third. In one sense a *harmonia* is the *logos* of a mixture, i.e., the ratio of the elements, which may be expressed mathematically. In another it is a combination (*synthesis*) of physical objects, probably in the sense of the arrangement of a number of physical parts but perhaps also as the complex of those parts in that arrangement. There appears also to be a fourth sense of *harmonia* that Aristotle ignores, in which a *harmonia* is something causally dependent on a certain disposition of materials; e.g., a melody is distinct from the strings that produce it, and equally from the tuning of the strings, though without strings there could be no tuning, and without tuning no melody. The word has this sense especially in musical contexts, meaning variously "scale," "mode," or generally "music" (see Liddell and Scott). Given, then, that the elements in question are those that compose the human body, the hot, the cold, etc., (which are presumably thought of as different kinds of stuff), there appear to be four possible interpretations of the thesis that the soul is a *harmonia* of those elements:

1. the soul is identical with the ratio or formula according to which the elements are combined to form the living man;
2. the soul is identical with the mixture or combination of those elements according to that formula;
3. the soul is some entity produced by the combination of those elements according to that formula, but distinct alike from them and from the formula itself;
4. the soul is identical with a state of the bodily elements, viz., the

state of being combined according to that formula.

It might be objected at this point that the third alternative is illusory, since even when the *harmonia* is a scale or melody it must be considered identical with a mixture of elements. This seems implausible on the assumption that the elements in question are strings or other physical objects composing the instrument that produces the music, but that assumption is mistaken. Just as the elements of a physical organism such as a living human body are the hot, the cold, etc., so the elements of a piece of music are the high and the low, which are conceived of as being mixed together in the proper proportions to give the right notes, either in the sense that each note is thought of as consisting of so much of the high mixed with so much of the low, or in the sense that each mode or scale is produced by combining so many high notes in fixed ratios with so many low notes. The elements, therefore, of a musical *harmonia* are themselves musical entities, the high and the low, not the physical objects that produce the sounds. This theory is clearly expressed for instance in the pseudo-Aristotelian treatise *De Mundo* 396b7 ff. (DK 22 B 10): "Music makes a single *harmonia* out of different sounds by mixing together high and low, long and short notes." On this view of musical *harmonia*, then, the *harmonia* cannot be separated from its elements, and so this view does not admit interpretation 3 as an alternative to the others.

But while this view of the nature of musical *harmonia* appears to have been the standard view of musical theory and gives the most exact parallel to other kinds of *harmonia* (e.g., the formation of physical substances out of the elements), it is emphatically not the view of musical *harmonia* that Simmias uses to illustrate his thesis. For his presentation of that thesis involves positing a parallelism between two relations, of each of which the terms are (a) a physical object and (b) a nonphysical entity causally dependent on that object. Thus corresponding to the incorporeal soul we have the musical *harmonia*, which is "invisible and incorporeal and all-beautiful and divine" (85e5-6), while corresponding to the physical body we have not the high and the low but the physical strings and pegs of the lyre, which can be broken apart and left lying around after the *harmonia* has vanished. It is true that Simmias slightly distorts the parallel when he says (85b5-c1) that the soul is a *harmonia* of the hot, cold, etc. in the body, since a more exact parallel to the strings, etc., of the lyre would be provided by the limbs and organs of the body than by their microscopic elements. But the essential point is the contrast of the incorporeal product with its physical cause, and for that contrast it

is unnecessary for Simmias to distinguish the macroscopic parts of the body from their own elements, which are no doubt conceived of as minute but equally corporeal parts. The relation of musical *harmonia* to its elements, which Simmias is using, cannot therefore be that between a scale or tune and the musical elements of high and low, etc., but must be that between a musical instrument and some nonphysical entity produced by a certain state of the instrument.

This enables us immediately to eliminate interpretation 2 above. For it would clearly be absurd to make a sharp contrast between the physical elements and the nonphysical *harmonia*, if the latter were just the elements themselves in a certain arrangement. One might as sensibly contrast the invisible, incorporeal plum pudding with the gross, earthy suet, raisins, flour, etc., that compose it. This still leaves three of the original four alternatives, that the relation of the soul to the body is (using the original numbering) (1) that of the ratio of the tuned strings to the strings themselves, or (3) that of the music produced by the instrument to the instrument itself, or (4) that of the state of being in tune to the strings.

There is no conclusive evidence from the text which alternative Plato had in mind, or indeed whether he had distinguished the three. Various phrases give some hints, but these are conflicting and inconclusive. Thus the description of musical *harmonia* as "all-beautiful and divine" might seem most readily applicable to the music produced by the instrument; but when we reflect that the speaker is a pupil of the Pythagorean Philolaos, who might therefore be expected to have a lively reverence for numbers as the source of all things, this argument seems to have little force as between alternatives 1 and 3. Rather stronger is the argument from Simmias' statement at 92d2 that the soul-*harmonia* doctrine is accepted by most people; surely, it may be argued, this indicates that the soul is something distinct from a mathematical ratio, since such an obscure theory can never have been held by the majority. On the other hand, the view that the soul is something nonphysical, which is yet dependent on a certain state of the body, so that when that state is disrupted the soul is dissipated, might seem to be quite congenial to common sense. But against this we have the comparison of the soul at 86c6-7 to "*harmoniai* in sounds and in all the works of the craftsmen." "All the works of the craftsmen" must include statuary and painting, and probably carpentry and house-building as well. Where, in the products of these arts, are we to look for the nonphysical product of the physical elements? Surely in the harmony or proportion of the constituent parts, as exemplified by the proportions between different amounts of different pigments,

or by the relations between the dimensions of the various parts of a statue or a piece of furniture. It would be too fantastic to suggest that to every well-made table there corresponds a nonphysical entity that is related to the disposition of its parts as the nonphysical soul is to the disposition of the bodily elements, or as the nonphysical music is to the disposition of the strings. This comparison, then, tends to support alternatives 1 and 4, rather than 3, which was suggested by the claim of popular acceptance for the *harmonia* thesis. Further difficulty is created by the description of the soul at 86b9 as a mixture (*krasis*) of the bodily elements. The word *krasis*, which is regularly used as a synonym for *harmonia* (e.g., Aristotle, *De Anima* 408a30-1), commonly occurs, like the English "mixture" in contexts that leave it open whether the word refers to the compound of elements that are mixed, or to the state of those elements of being mixed. We have seen that the former alternative is unacceptable, but what about the latter? Can Plato mean that the soul is identical neither with a ratio nor with any nonphysical product of a ratio, but rather with a certain state of the body, viz., the state in which the elements of the body are in a certain ratio? While on the one hand this would give a fair account of the comparison of the soul with *harmoniai* in works of art, on the other hand it fits rather ill with the sharp contrast between the invisible, divine musical *harmonia* and the physical instrument, while again it might well seem very dubious that most people believe that the soul is nothing other than a bodily state. There appear, then, to be hints of support in the dialogue for all three possible interpretations of the soul-*harmonia* thesis, which might suggest that Plato has failed to distinguish these alternatives. Before leaving this question, however, we should look at some evidence from other sources, to see whether these throw any light on Plato's meaning.

First, there is the fact already mentioned that in *De Anima* 1.4 Aristotle ignores the possibility that on the *harmonia* theory the soul might be a nonphysical entity causally dependent on the ratio of bodily elements, while explicitly mentioning the possibilities of its being identical with that ratio and of its being identical with an arrangement of parts, which is itself ambiguous between "identical with the parts as arranged" and "identical with the state of being arranged which characterizes the parts." Not only does he not give the first-mentioned possibility as a possible interpretation of the thesis, but he appears to introduce it as an alternative view of the soul, which would be unacceptable to an adherent of the *harmonia* theory. After producing objections first to the suggestion that the soul is a combination of limbs and then to the suggestion that it is the

ratio of the mixture of elements, he adds (408a20-1) "Is it the ratio which is the soul, or is it rather something separate which comes to be in the parts of the body?" The implication is that he is here suggesting a more plausible alternative theory, not giving an interpretation that a supporter of the theory would accept as expressing his meaning. Yet in order to take this as a basis for the interpretation of the thesis of the *Phaedo* we should have to have some reason to believe that Aristotle's target in the *De Anima* is specifically the thesis proposed in that dialogue, rather than any other version. Such reason is lacking. Aristotle describes the theory as widely current (407b27-8), but does not ascribe it to anyone in particular, and does not mention the *Phaedo* in his discussion. We know that a version of the theory was held by Aristotle's follower Aristoxenus (for the evidence see H. B. Gottschalk, "Soul as *Harmonia*," *Phronesis* 16 (1971):179-98), and it is possible that a version different from that of the *Phaedo* was maintained by Philolaos (see below). The *harmonia* theory was, then, a current theory of the soul,<sup>2</sup> and we are safest to suppose Aristotle to be attempting to expose the fundamental errors of the theory as such, rather than give an exact exposition of any version of it.

One might hope to throw some light on the precise sense of the theory by considering its origins, but here too it is impossible to reach any positive conclusions. None of the speakers in the dialogue attributes it to any named philosopher, but since Simmias says that "we" hold the soul to be a *harmonia* (88b6-7) and Echecrates that he has always been very impressed by that thesis (88d3-4), it would be natural to assume that it was current in the Pythagorean circle to which they belonged. Though they are described as pupils of Philolaos (61d-e, cf. Diogenes Laertius VIII-46), the theory itself is not ascribed to Philolaos by any writer earlier than Macrobius (fourth to fifth centuries A.D.), who says that Pythagoras and Philolaos held that the soul is a *harmonia* (DK 44 A 23). It is not clear how much reliance can be put on this testimony, since there is obviously a possibility that it may derive ultimately from this very passage of the *Phaedo*. But whatever the truth about that, Philolaos' view of the soul cannot be reconciled with the *harmonia* theory as expounded by Simmias. For at 61a-62b it is implied that Philolaos taught that suicide is wrong on the ground that the soul is put by the gods in the body as a prison for a set time, and must not seek to escape before the time of its release, though a philosopher will welcome death, presumably because his soul will have a better existence in separation from the body. This is supported by a quotation from Philolaos given by Clement of Alexandria (DK 44 B 14): "The soul is yoked to the body and as it were buried in

this tomb as a punishment." The conclusion is plain that unlike his pupils who take part in the dialogue Philolaos believed that the soul exists independently of the body. It is not impossible that he may have held some version of the theory, in which the soul was a nonphysical entity whose association with the body depended on the maintenance of the proper bodily ratio, but the divergence from the view expressed by Simmias is so great that it is obviously fruitless to attempt to interpret the latter in such a way as to assimilate it to some conjectural reconstruction of Philolaos' views.<sup>3</sup>

I conclude, then, that not only is there no evidence that the soul-*harmonia* thesis definitely identifies the soul either with a ratio of bodily elements or with the state of being in that ratio or with some entity dependent on the possession of that ratio, but that we can best account for what is said on the assumption that Plato did not clearly distinguish the three possibilities. In considering the arguments against the thesis we shall therefore regard them as concerned with a thesis containing those three alternatives in undifferentiated form.

Socrates' first counterargument requires little comment. He points out that the thesis is inconsistent with the doctrine accepted earlier that all knowledge is in fact recollection of what the soul had learned in a previous, disembodied existence. No *harmonia* can exist unless the elements of which it is a *harmonia* are already in existence and hence if the soul is a *harmonia* of the bodily elements it cannot have had a previous nonbodily existence (91e-92e). This argument is cogent against any interpretation of the thesis; obviously a bodily state cannot exist unless some body exists of which it is a state, and equally obviously a non-physical entity causally dependent on a ratio of bodily elements cannot exist before those elements have been combined in that ratio. A defender of the thesis might, however, argue that it is not cogent against the identification of the soul with the mathematical ratio itself. For a ratio, being a timeless mathematical entity, cannot itself be said to come into existence whenever it is embodied in some particular material. Since it exists equally at all times, it may truly be said to have existed before a certain body came into being, and hence the argument from recollection does not refute this version of the thesis. This defense is not, indeed, adopted by Simmias, who agrees that his thesis is inconsistent with the doctrine that knowledge is recollection. Nor is it difficult to see why. For it is possible to use this defense only at the cost of making the soul-*harmonia* a universal; if a pair of elements combines in the ratio 3/4, then indeed that ratio existed before the combination of those elements, but what existed

was *the* ratio 3/4, i.e., the very same ratio that is exemplified whenever three units are related to four units. Thus anyone holding this theory would have to admit that many things must have the same soul, including things generally reckoned inanimate, e.g., geometrical diagrams, since the same ratio that is embodied in a human being and is his soul may also hold between certain lines and angles. It is not, of course, impossible that anyone might have believed something like this; it might, for instance, provide a theory to account for transmigration. Empedocles would on this view have been a bush, a fish, etc. (DK 31 B 117) because one and the same ratio would have been embodied in bush, fish, and Empedocles; i.e., they all had the same soul. Simmias, however, will have none of this; if his version of the theory is interpreted as making the soul a mathematical entity, it must be such an entity individuated by being embodied in these bodily elements. As such it cannot exist independently of the elements by reference to which it is individuated, any more than Socrates' height can exist independently of Socrates, though in the sense in which Socrates' height is a universal, say 5 feet 6 inches, that length may be said always to have existed whether or not Socrates exists. This way of looking at the thesis has the advantage of preserving as a necessary truth the thesis that different persons have numerically different souls, whereas on the other interpretation it might be discovered as the result of physiological investigation that two different people had the same soul. It leaves the thesis open, however, to attack on the grounds of inconsistency with the doctrine that knowledge is recollection; whether one considers it refuted on that ground will naturally depend on the strength of one's conviction of the truth of that doctrine.

The remaining arguments are more problematical, in that commentators have disagreed not merely on their conclusiveness but also on how many arguments are employed, and precisely what these arguments are. Like Hicken, Bluck (R. S. Bluck, *Plato's Phaedo* [London: Routledge & Kegan-Paul, 1955]), and Gallop, I discern two arguments, as opposed to the four specified by Philoponus in his commentary on *De Anima* 1.4.<sup>4</sup> These arguments are not, however, presented consecutively; at 92e4-93a10 Socrates gives a set of propositions A1-A4 that are not immediately used in the argument.<sup>5</sup> Instead at 93a11-12 he begins a new argument by formulating a principle B1 that is to some extent independent of A1-A4. This argument (argument *B*) continues to its conclusion at 94a12-b3. Then at 94b4 Socrates returns to A1-A4, which he uses to construct the second argument (argument *A*), whose conclusion is reached at 95a2. While I shall deal first with argument *B*, it is necessary

first to look at A1-A4 in order to determine their relation to B1.

Socrates begins by securing Simmias' acceptance of the proposition that the properties of a *harmonia* are determined by those of its elements (92e4-93a2; A1). We then have three successive applications of this principle, first to all activities and passivities of the *harmonia* (a4-5; A2) and then to some particular activities and passivities that are excluded by the principle. It is impossible for a *harmonia* to lead or control its elements, but it must rather be controlled by them (a6-7; A3), and it is impossible for it to be affected or behave in any way contrary to that which its elements determine (93a8-9; A4). At 93a11-12 we have the principle that marks the beginning of argument *B*: "Well now, doesn't every *harmonia* have to be the kind of *harmonia* that corresponds to the way that it is ordered?" (B1).<sup>6</sup> It is not easy to find a translation that is both exact and comprehensible but the next sentence, giving an application of the principle, makes the meaning fairly clear; if a *harmonia* is more ordered it is more (of) a *harmonia*, and if it is less ordered it is less (of) a *harmonia* (93a14-b2: B2). The sense of B1 itself can then best be expressed formally, as follows, that where *F* stands for an adjective that can apply to a *harmonia*, and where *F*-ly stands for the adverb formed from the adjective for which *F* stands, then for all *X*, if *X* is a *harmonia*, if *X* is ordered *F*-ly, *X* is an *F* *harmonia*. While this certainly goes beyond anything that is said in A1-A4, this "formal" account of the dependence of a *harmonia* on what gives rise to it may nonetheless be seen as continuing the line of thought begun there. The crucial difference, emphasized by Gallop, *Plato, Phaedo*, 158, is that whereas in A1-A4 we are concerned with the dependence of the *harmonia* on its elements, B1 states the dependence of the *harmonia* on the order or arrangement of the elements.

Argument *B* proceeds by way of three further premises, B3 that a soul is no more or less (of) a soul than any other (b4-6), B4 that there are some good souls and some bad (b8-c2) and B5 that a good soul is in order and a bad soul out of order (c3-10). None of these premises is felt to require any justification or explanation; the sense of the third is clearly that the good man is not a prey to the conflicting desires and impulses that are the mark of the bad man, but has all his wants properly under control with a view to the attainment of the right ends. We now come to one of the most problematical passages in the argument: at d1-5 Socrates says that premise B3 is the same as the proposition (B7) that no *harmonia* is more or less (of) a *harmonia* than any other, and Simmias agrees. Of course B3 is not as it stands equivalent to B7, and the question is what additional assumptions Plato must have used in order to produce what he

regarded as a valid equivalence. Clearly we cannot derive such an equivalence simply by making the most obvious assumption, viz., the assumption under examination in this argument, that the soul is a *harmonia* (B6), since taken together with B3 that would still allow that some *harmoniai* might be more or less *harmoniai* than others. But did Plato see that? I am inclined to think that he did not, but rather, assuming the soul to be a *harmonia*, took this to imply that whatever is true of soul is also true of *harmonia* (using the terms in the unquantified style familiar from Aristotle). In effect this is to confuse implication with equivalence, which seems a possible error for Plato to commit at this stage in his philosophical development. (See Gallop, *Plato*, Phaedo, 162-63.)

The standard modern interpretation of this sentence, adopted by Archer-Hind, Bluck, Hackforth, and Hicken (but not by John Bumet, *Plato's Phaedo* [Oxford: Oxford University Press, 1925] or Gallop) differs from the above in taking Socrates to be asserting not a general proposition about all *harmoniai* but a specific proposition about the sort of *harmonia* that souls are, viz., that no soul-*harmonia* is more or less a *harmonia* than any other. As this requires an admittedly unnatural reading of the text as it stands, many scholars (see Hackforth's note, *Plato's Phaedo*, 116) have suggested deleting the word *harmonias* from d4, thus making the sentence read "And this (namely, the admission that no soul is more or less a soul than any other) is the admission that no (soul) is more or less a *harmonia* than any other." But since this emendation lacks any manuscript authority and destroys what looks like a very emphatic and deliberate parallelism of sentence construction, it is worth asking whether there are cogent grounds for emending the text or for reading the received text in a sense other than its natural one. The strongest ground appears to be that urged by Hicken, that since the argument is to depend on the assumption that some *harmoniai* (in particular, goodness) admit of degrees, it would be flatly inconsistent if Plato also relied on the assumption that no *harmonia* admits of degrees. I doubt the cogency of this argument, which seems to depend on a confusion in the notion of "degrees of *harmonia*." For the thesis that some *harmoniai* (e.g., goodness) admit of degrees comes to this, that some things, e.g., the parts of the soul, may be so arranged as to approximate more or less closely to some norm that represents the perfect arrangement of those things. But that is in no way incompatible with the thesis that I take Plato to be asserting at 93d1-5, viz., that if what a thing is is a *harmonia*, it cannot be more or less a *harmonia* than anything else that is a *harmonia*. This amounts to an extension of the truism "Everything is what it is," and applies alike to

perfect and to imperfect orderings. Every ordering of parts of the soul, at whatever remove from the norm, is equally an ordering of parts of the soul. There is, then, no general inconsistency between the theses "No *harmonia* is more or less a *harmonia* than any other" and "Some things are more ordered (in Platonic terms 'partake more of order') than others." Plato, however, thinks that contradiction is generated if one says that one *harmonia* is more ordered than another; that he is wrong even in this restricted thesis will be seen once the argument is viewed as a whole.

The next step (93d6-8; B8) is that something that is neither more nor less a *harmonia* is neither more nor less ordered; this follows directly by contraposition from B2. Another problematic sentence follows (d9-11; B9): "And does that which is neither more nor less ordered partake more or less of order, or to just the same extent? To the same extent." At first sight it might appear that this is the converse of B8. But, firstly, in contrast to the previous sentence, where the subject is "that which is neither more nor less a *harmonia*" the predicate of B9 is "partakes of (i.e., is characterized by) *harmonia* more or less." While one might indeed see here a confusion between predication and identity, it is more charitable to take the shift in terminology as indicating that a new point is being made by the introduction of a further premise. Secondly, if B9 is read as the converse of B8, it has no subsequent role in the argument; whereas if it is read as "Something which is neither more nor less ordered has neither more nor less order," we have a straightforward argument, as will be seen immediately.

Socrates next concludes (d12-e2; B10) that no soul is more or less ordered than any other, giving as premise B3. In fact B10 follows, not from B3, but from B6, 7, and 8 (see appendix on page 65). Socrates' derivation of B10 from B3 presumably indicates that he is relying on the fallacious derivation of B7 from B3 (together with B6). From this point the argument is straightforward. From B9 and B10 it follows that no soul is more or less ordered than any other (e4-5; B11), and hence by B5 and B11 that no soul is better or worse than any other (e7-94a10; B12). It is agreed (94a12-b3) that this conclusion is absurd, and hence one of the premises from which it is derived must be false; obviously, the premise to be rejected is the assumption that the soul is a *harmonia*.

It appears, therefore, that we have in argument *B* a single argument that is, despite some obscurities, clear in its main lines and (perhaps not so clearly) fallacious. The flaw is not simply the fallacious equivalence of B3 and B7, since one might patch this up by introducing B7 as a premise; it is perfectly plausible to suggest that, where *F* is a predicate

saying what kind of thing its subject is, if *A* and *B* are both *Fs*, *A* can't be more (of) an *F* than *B*. That emendation still leaves a fallacious argument, though the text leaves room for more than one account of the fallacy.

One possibility is the following. The argument is invalid through Plato's failure to recognize that ". . . is a *harmonia* (ordering)" and ". . . is ordered" are incomplete predicates, requiring to be completed as ". . . is an ordering of elements of type *E*" (hereafter ". . . is an *E*-ordering") and ". . . is ordered with respect to elements of type *E*" (hereafter ". . . is *E*-ordered"). The distinction is crucial, since one and the same thing may consist of elements of two types *E* and *E'*, such that that thing is an ordering of elements of type *E*, while it neither is an ordering of elements of type *E'* nor is characterized by order in respect of elements of that type. Imagine a university composed of independent multidisciplinary colleges, which also has a faculty structure crossing collegiate boundaries. Imagine further that, while the relations of the faculties to one another and to the university are fully organized, intercollegiate relations and relations between the university and the colleges are anarchic. (Readers familiar with the Universities of Oxford and Cambridge will note elements both of verisimilitude and of exaggeration.) The questions "Is the university an organization?" (I treat the term *organization* as interchangeable for the purpose of this illustration with *ordering*) and "Is the university organized?" (= "ordered") have no determinate answers. It is an organization of faculties, but (despite being composed of colleges) it is not an organization of colleges, nor is it organized in respect of the colleges. The situation of the soul, on the assumptions of the *harmonia* theory, is parallel. It is *ex hypothesi* an ordering of the bodily elements (allowing for the ambiguity of that thesis as discussed above). In virtue of that relation of bodily elements the soul consists of psychic elements, desires, intellect, etc., which may themselves be organized in a coherent way (the state of a good soul), or may lack organization (the state of a bad soul). No soul, good or bad, is more an ordering of bodily elements than any other, for an ordering of bodily elements is just what a soul is. But orderings of bodily elements may be more or less ordered in respect of psychic components.

The various steps of the argument have now to be rephrased in terms of the complete predicates ". . . is an *E*-ordering," etc. The crucial changes are as follows:

for B2 substitute B2'—Any *E*-ordering that is more *E*-ordered is more

of an *E*-ordering, and any that is less *E*-ordered is less of an *E*-ordering;  
 for B5 substitute B5'—A good soul is *ps*-ordered (possesses psychic order), a bad soul is *ps*-disordered (lacks psychic order);  
 for B6 substitute B6'—The soul is a *ph*-ordering (i.e., an ordering of physical elements);  
 for B7 substitute B7'—No *E*-ordering is more an *E*-ordering than any other;  
 for B8 substitute B8'—That which is neither more nor less an *E*-ordering is neither more nor less *E*-ordered;  
 for B9 substitute B9'—That which is neither more nor less *E*-ordered has neither more nor less *E*-order.

It is now clear that the argument fails to lead to a *reductio*. From B6', 7', and 8' we derive B10': "No soul is more or less *ph*-ordered than any other." And from B9' and B10' we derive B11': "No soul has more or less *ph*-order than any other." But from B5' and B11' it is impossible to derive B12. For by B5' bad souls possess less *ps*-order than good souls, while by B11' all souls have equal amounts of *ph*-order. B12 does not follow, and hence there is no contradiction with B4 and no *reductio*. It is, then, not inconsistent to maintain that an entity that arises from the organization of bodily elements may itself contain parts or elements of another sort that lack organization, or alternatively that a certain organizational state of bodily elements (i.e., the state of being ensouled) may be in a particular case further characterizable as a state of psychic disorganization.

An alternative diagnosis of the fallacy is provided by Gallop, *Plato, Phaedo*, 163-66. On this view the flaw is not the failure to supply different completions for an incomplete predicate at different stages of the argument but rather a failure to observe an ambiguity in the term *harmonia* between "tuning" or "order" ("attunement<sub>1</sub>" in Gallop's terminology) and "correct tuning" or "good order" ("attunement<sub>2</sub>"). While every attunement<sub>1</sub> is equally an attunement<sub>1</sub> different attunements<sub>1</sub> may be characterized by attunement<sub>2</sub> to different extents. (Essentially the same point is made by pointing out that the sentence "*A* is more ordered than *B*" is ambiguous between "*A* is more of an ordering than *B*" and "*A* is better ordered than *B*.") Hence B12, "No soul is better or worse than any other," does not follow from B11: "No soul has more or less order than any other." For the predicate in B12 is an instance of ". . . is characterized by attunement<sub>2</sub>" whereas that in B11 is ". . . is characterized by attunement<sub>1</sub>." As far as I can see, this suggestion fits the text as closely as

that given above. Since Plato is arguing from the premise that all souls share equally in the ordering of physical elements to the (absurd) conclusion that all share equally in the good ordering of psychic elements, it is not surprising that the text should leave it open whether the failure of the argument turns on the slip from “*ph*-order” to “*ps*-order” (my suggestion) or on that from “order” to “correct order” (Gallop’s suggestion). Either is sufficient to generate a fallacy, while the text allows for both.

Argument *A* is resumed immediately. It is agreed that in a sensible man the soul controls and opposes bodily inclinations such as hunger and thirst (94b4-c1; A5). Socrates now (c3-c6) recalls steps A3 and A4, to the effect that a *harmonia* can never control or be opposed to its elements. Hence (94e8-95a2) the soul cannot be a *harmonia*. This concludes the discussion of the thesis.

In this case it is clear that the crucial slip is that from *ph*-order to *ps*-order. (This may be some small reason for preferring that account of the fallacy in argument *B*.) If the *harmonia* thesis is to be refuted by this argument, the soul’s control of bodily inclinations must be a case of what is denied by A3, viz., the soul’s controlling its elements. But according to the thesis, the elements of which the soul is a *harmonia* are not any sort of inclinations but bodily elements, the hot, the cold, the wet, the dry, etc. On Simmias’ statement of the thesis the activity of the soul is determined by the interrelations of those elements, so that, for instance, a certain mixture of hot and cold will produce anger in the soul, or a certain proportion of dry to wet the desire for a drink. But in argument *A* Socrates treats such events as being angry or wanting a drink as themselves impulses of the physical elements of which the soul is *ex hypothesi* a *harmonia*, and insists on the incompatibility of A3 and A4 with the view that the soul opposes these impulses, in the sense that the reason often opposes such desires. Clearly, there is no inconsistency. All that a defender of the *harmonia* thesis need say is that in the case of such a conflict of reason and desire we see, not the soul-*harmonia* opposing and controlling its elements, but rather one part of the soul-*harmonia* opposing and controlling another. And he might add that of course the controlling part is as determined as the controlled part by some disposition of bodily elements. In effect this would be to replace the soul-body dualism of the *Phaedo* with an account akin to that of the divided soul in the *Republic* and *Phaedrus*, with the addition of a thesis of physicalistic determinism of the functioning of all parts of the soul.

Plato’s intention in arguing against the *harmonia* thesis is no doubt partly to resist this determinism, in support for the insistence on the

autonomy of the rational soul that pervades the *Phaedo*. Yet our discussion of argument *B* has shown that it is hard to disentangle arguments against the thesis that the soul is a *harmonia* of physical elements from the thesis that it is a *harmonia* of elements of any kind. (Gallop’s analysis of the fallacy is independent of the nature of the elements.) Yet the development of Plato’s thought requires that he should at some stage have made that distinction, since there seems to be a perfectly good sense in which the tripartite soul of the *Republic* and *Phaedrus* may be called a *harmonia*, in that it is a composite entity composed of parts whose relations affect its functioning as a whole. Perhaps argument *B* of the *Phaedo* discussion was directed specifically against the account of the soul as a *physical harmonia*, and was not intended to apply to any other sort. (This would count against Gallop’s interpretation.) Alternatively, assuming Gallop to be right, Plato had spotted the fallacy by the time he developed the theory of the tripartite soul and saw that that theory was safe against the arguments of the *Phaedo*. It is, however, necessary to attribute extraordinary obtuseness to Plato if one accepts, with Hicken and others, that the arguments of the *Phaedo* are conclusive against the thesis. For if those arguments are sound, they refute the theory of the tripartite soul, in which case the whole political organization of the *Republic* is based on a psychological theory that Plato had already (assuming the priority of the *Phaedo*) refuted.

### Appendix: Analysis of the Arguments

92e4-	<b>A1</b>	The properties of an ordering are determined by those of its elements.	P(remise)
93a2			
a4-5	<b>A2</b>	The activities and passivities of an ordering are determined by those of its elements.	from A1
a6-7	<b>A3</b>	An ordering cannot control its elements but must be controlled by them.	from A2
a8-9	<b>A4</b>	An ordering cannot be affected or behave in any way opposed to (the behavior of) its elements.	from A2
a11-12	<b>B1</b>	Any ordering that is ordered <i>F</i> -ly is an <i>F</i> ordering.	P
a14-b2	<b>B2</b>	Any ordering that is more ordered is more of an ordering, and any that is less ordered is less of an ordering.	from B1
b4-6	<b>B3</b>	No soul is more or less of a soul than any other.	P

b8-c2	<b>B4</b>	There are some good souls and some bad.	P
c3-10	<b>B5</b>	A good soul is ordered (possesses order); a bad soul is disordered (lacks order).	P
(under- stood)	<b>B6</b>	The soul is an ordering.	P
d1-5	<b>B7</b>	No ordering is more of an ordering than any other.	from B3, B6 (invalid)
d6-8	<b>B8</b>	That which is neither more nor less an ordering is neither more nor less ordered.	from B2
d9-11	<b>B9</b>	That which is neither more nor less ordered has neither more nor less order.	P
d12-e2	<b>B10</b>	No soul is more nor less ordered than any other.	from B6, B7, B8
e4-5	<b>B11</b>	No soul has more nor less order than any other.	from B9, B10
e7-94a10	<b>B12</b>	No soul is better or worse than any other.	from B5, B11
a12-b3	<b>Conc.</b>	Since B12 contracts B4, B6 is false.	from B4, B6, B12 (RAA)
b4-c1	<b>A5</b>	The soul controls and opposes bodily desires.	P
(under- stood)	<b>A6</b>	An ordering opposes and controls its elements.	from B6, A5
c3-e6	<b>A7</b>	A6 contradicts A3 and A4.	from A3, A4, A6
e8-95a2	<b>Conc.</b>	B6 is false.	from B6, A7 (RAA)

Macmillan, 1883); and R. Hackforth, *Plato's Phaedo* (Cambridge: Cambridge University Press, 1955), are cogently stated by W. F. Hicken, "Phaedo 93a11-94b3," *Classical Quarterly* 48 (1954): 17-18, and by Gallop, *Plato, Phaedo*, 160-61.

5. See appendix.

6. The verb translated "order" (*harmozein*) may equally properly be rendered "attune," "arrange," or "organize." I shall adhere to "order" throughout.

## Notes

1. An earlier version of this paper was written in 1970 and delivered (in the author's absence) at the meeting of the Eastern division of the American Philosophical Association in December of that year (see *Journal of Philosophy* 67 (1970): 724). It is that version which is mentioned by David Gallop, *Plato, Phaedo* (Oxford: Clarendon, 1975), 157. In preparing this revision I have profited not only from Gallop's valuable discussion but also from the comments of Gareth Matthews on the 1970 version of the paper.

2. Lucretius discusses the thesis that the "sensus animi" is a *harmonia* at III.98-135, attributing it merely to "(the) Greeks."

3. For a fuller discussion (leading to similar conclusions) see Gottschalk, "Soul as *Harmonia*," *Phronesis* 16 (1971): 179-98.

4. The grounds for rejecting Philoponus' analysis, which is followed, not without incoherence, by R. D. Archer-Hind, *The Phaedo of Plato* (London:

## Supervenience and the Thesis That the Soul Is a *Harmonia*

Ellen Wagner

At *Phaedo* 85e-86d Simmias presents his thesis that the soul is a kind of *ἀρμονία* (*harmonia*) or attunement of bodily elements: the hot and cold, moist and dry. The *harmonia* is intended as a counterexample to the just-elaborated Argument from Affinity, which had claimed that the soul is immaterial, incomposite, and indestructible, and therefore immortal. Commentators have reviewed in detail both this thesis and Socrates' counterarguments to it at 92a-95a.<sup>1</sup> What they have not done, however, is to explain why Socrates considers Simmias' thesis so powerful that he presents three quite sophisticated arguments to defeat it: why does it merit so much attention? The most convincing answer, I will argue, is that Socrates knows the Argument from Affinity has no defense against the *harmonia* thesis, in the interpretation I offer here. A *harmonia* possesses a number of the properties Socrates attributes to the soul in the Affinity Argument, yet it is clearly mortal. Even if we accept all of the fundamental claims made by Socrates on behalf of the soul in the argument, Socrates fails to establish that the soul is immortal.

The supervenient dualist interpretation I present makes explicit the way Simmias' objection threatens the Platonic substance dualism assumed in the *Phaedo*'s arguments for immortality.<sup>2</sup> On my view, the soul is an entity distinct from the body by virtue of two non-overlapping sets of properties, exactly as the Affinity Argument stipulates. However, the *harmonia* is a successful counterexample to

that argument because, although the supervenient soul possesses certain of the properties of the immortal kind of being, its causal dependency upon the elements of the material body renders it ineluctably mortal. In this way, the *harmonia* thesis shows that the Affinity Argument fails to establish that the soul belongs to the kind of existents constituted by the Forms. As I will show, my interpretation of Simmias' thesis explains why the thesis provokes Socrates to put forward three ingenious counterarguments, and also provides a compelling diagnosis of what goes awry with the longest, most complex of these arguments. It will be helpful first, however, to revisit the argument Simmias takes himself to be refuting in his thesis: the argument that the soul is immortal because of its resemblance to—or affinity with—the Forms.

### 1. The Argument from Affinity

Simmias at this point in the *Phaedo* is unconvinced by the latest argument for immortality Socrates has presented: the Argument from Affinity (78b-80b), which receives its name from its fundamentally analogical nature.<sup>3</sup> Socrates proceeds from the premise that there are two classes of existents to which the soul might belong (79a6): “would you like us to posit two kinds of beings?” (δύο εἴδη τῶν ὄντων).<sup>4</sup> One kind consists in the Good, the Beautiful, the Equals, the Just, and their like (74c5, 75c11-d1, 76a1-9, 76d6-9), while the second is composed of such things as equal sticks and stones (74b4-9), the many beautiful men, cloaks, horses, and things of that kind (78d10-e1) and all other things homonymous with the objects of the first kind. The former are invisible and by extension imperceptible by any sense, constant, noncomposite, indissoluble, and indestructible. The latter possess all of the contradictories of these properties, being visible, changing, composite, dissoluble, and therefore liable to be destroyed. In three distinct subsections of argumentation Socrates concludes that the soul is more similar (ὅμοιότερον) to what is invisible (79b16), what is constant (79e4-5), and what is divine (80a8) than is the body. Surprisingly, he directly attributes to the soul only one property—constancy—and then to an extremely limited degree: it is constant and unvarying in the presence of the constant kind of being (περὶ ἐκεῖνα ἀεὶ κατὰ ταῦτα ὡσαύτως ἔχει, 79d5). As in any argument from analogy, Socrates concludes from the fact that the soul is similar to the Forms in several respects that it shares a further property with them: just as the Forms are *completely* indissoluble, so

is the soul *nearly* so.<sup>5</sup> He does not stop with indissolubility, however. In further discussion, Socrates notes that even if the body decays, certain parts of it such as the bones and sinews remain for a long time; these parts are—the general decay of the body notwithstanding—so to speak, immortal (ὅμως ὡς ἔπος εἰπεῖν ἀθάνατά ἔστιν, 80d2). So the soul, properly purified by the correct philosophical practice, will in truth spend the rest of time among the gods (ώς ἀληθῶς τὸν λοιπὸν χρόνον μετὰ θεῶν διάγουσα, 81a9). In other words, we may expect the soul to be immortal.

That Socrates has doubts about the success of his argument is plain from the qualifications of the three preliminary conclusions noted above, as well as from the language of his final conclusion at 80b9-10: the soul is indissoluble or, at any rate, *something close to it*—ψυχῆ δὲ αὐτὸν τὸ παράπαν ἀδιαλύτῳ εἶναι η̄ ἐγγύς τι τούτου. The soul's definite membership in the kind of the Forms is not unconditionally established. Clearly a strong counterexample would quickly undermine the argument, perhaps fatally; and it is just this counterexample Simmias believes he can offer as the thesis that the soul is a *harmonia*.

### 2. The Relevant Text

At 85e3-86b5 Simmias presents his alternative model of the soul: the soul is a *harmonia* of bodily elements, “the hot and cold, dry and wet, and such things” (86b8-9), like the *harmonia* of a lyre and its strings. The conclusion of the Argument from Affinity had previously attributed to the soul the character of being “most similar to what is divine, immortal, intelligible, uniform, indissoluble, unvarying and constant in relation to itself” (80b1-2). Similarly, an attunement, he claims, is something “unseen and incorporeal and very beautiful and divine in the tuned lyre” (ἀόρατον καὶ ἀσώματον καὶ πάγκαλόν τι καὶ θεῖόν ἔστιν ἐν τῇ ἡρμοσμένῃ λύρᾳ—85e5-86a1).<sup>6</sup> In opposition, the lyre and its strings are “corporeal bodies and composite (σύνθετα) and earthy and akin to the mortal” (86a2), just as the body had been characterized previously as “mortal, multiform, non-intelligible, dissoluble and never constant in relation to itself” (80b3-4). What is *prima facie* threatening to the argument about Simmias' new thesis is that a *harmonia* is most certainly composite: it is invisible and divine but not necessarily also simple.<sup>7</sup> In fact, the *harmonia* possesses all of the properties listed at 80b with the exception of noncompositeness.<sup>8</sup> More importantly, *harmoniai* in general may

have many of the properties of the immortal kind of being while they fail to be immortal.

Simmias' argument is as follows: even if the lyre and its strings are smashed and broken, the earthly materials persist for some time before they completely rot away. The attunement produced by them has properties which should, by the line of reasoning employed in the Argument from Affinity, make it immortal and so entail that it far outlast the material remains of the lyre. Yet clearly the attunement perishes immediately when the lyre is destroyed. Why should we not believe that the soul is an immaterial entity produced analogously by the material body and which perishes at bodily death? Thus, Simmias reasons, Socrates' argument is refuted.

We ought to notice first in Simmias' argument that the attunement is said to be *in* the tuned lyre (ἐν τῇ ἡρμοσμένῃ λύρᾳ—86a1). The use of the dative here indicates that the ἀρμονία is not merely produced by, or does not simply come out of, the lyre, but is in some way immanent.<sup>9</sup> This seems *prima facie* reasonable: where else could an attunement be present if not in the strings which now are tightened to specific tensions of the tuning appropriate for the mode of the music? If so, we will need to distinguish the *harmonia* from the lyre and define the specific relation of which they are the relata. Further, from the text it is difficult to determine the precise nature of the *harmonia*: is it a material or nonmaterial entity; is it identical to its components or something over and above them? Answering these questions will be essential to deciding just how worried Socrates ought to be about Simmias' counterexample.

### 3. The Nature of the *Harmonia*

Commentators in the literature have suggested several different interpretations of the thesis that the soul is a *harmonia*.<sup>10</sup> It might be materialist or nonmaterialist, given the ambiguity of the usage of *harmonia* in the passage, and hence three possible interpretations of the term might apply here:

- (1) The soul is the proportion itself in which the elements are combined to form the human being.
- (2) The soul is the combination of those elements in that proportion.

- (3) The soul is some entity produced by the combination of those elements in that proportion, but distinct both from them and from the proportion itself.<sup>11</sup>

One of these three possibilities may be eliminated fairly quickly. The second alternative is implausible on grounds both of the text and of the general argumentative strategy. Only as a nonmaterialist model of the soul is the attunement relevantly similar to the conception of the soul in previous arguments for immortality in the *Phaedo* (especially the Affinity Argument). The text itself shows that Simmias intends the thesis to be nonmaterialist: he is setting up a parallel relation between two pairs of entities, each of which consists in a material entity with its nonmaterial, causally dependent correlate. While the lyre and its strings are visible, corporeal, and mortal, the *harmonia* has the same nature as, and kinship with, what is divine and immortal (τοῦ θείου τε καὶ ἀθανάτου ὁμοφυῆ τε καὶ συγγενῆ, 86b1-2). The general grounds of the discussion presuppose that the soul is nonmaterial, as well: in his initial statement of his objection, Simmias has explicitly said that the *harmonia* in the lyre is unseen and immaterial (ἀόρατον καὶ ἀσώματον, 85e5). In the Argument from Affinity, invisibility is first of all the properties attributed to the soul; and Socrates means to ally the soul with the things accessible only by reason, in opposition not only to what is seen but to what is sensed in general. He asks Cebes, in fact, to contrast the constant kind of being with “those things you could touch and see and sense with the other senses” (Οὐκοῦν τούτων μὲν κανὸν ἀψαιο κανὸν ἵδοις κανὸν ταῖς ἄλλαις αἰσθήσεσιν αἰσθοιο, 79a1-2).<sup>12</sup> It therefore seems unlikely that Socrates would have been willing to accept as fair (δίκαια, 86d6) the grounds on which Simmias argues if Simmias had proposed that attunements were both divine and immortal, yet material.

Deciding between the two remaining interpretations is difficult, as both are nonmaterialist views. Given that Simmias is a student of the Pythagorean Philolaus, alternative (1), that the soul is a proportion, is appealing. But there is no determinate evidence that the “we” referred to (ὑπολαμβάνομεν) at 86b6-7 is anything other than the group of conversants gathered around Socrates, nor is there a monolithic Pythagorean view presented or assumed in the *Phaedo* to which Simmias can be alluding.<sup>13</sup> Such an orthodox view would have clarified the consideration of (1); in its absence we shall have to read more closely.

There are two possible ways to read (1). At 85e5 Simmias refers to the attunement in the tuned lyre, but he does not make clear

whether this attunement is the particular set of intervals at which the strings of any lyre may be tuned (as, for instance, those required for a melody in the dorian or phrygian mode) or that set of intervals as it occurs in one lyre's strings at that moment, whether tuned perfectly or inaccurately by the player. The first alternative, which I shall call (1a), is that the proportion exists as a universal. Thus, the soul is just the ratio of bodily elements (here we can imagine that ratio to specify the amounts of earth, air, fire, and water to compose the different kinds of components of the body, such as flesh and bone). The ratio would necessarily preexist any particular instantiation in order for that mass of matter to assemble itself correctly according to that proportion.<sup>14</sup> If so, the *harmonia* is best characterized as the universal proportion, which would be in effect the Form of soul.

This interpretation is not well supported by the text, which considers souls only as particulars belonging to persons with individual histories and moral states. It is the souls of Cebes, Simmias, and above all Socrates, that are the objects of concern in the *Phaedo*. Nowhere does Socrates refer to a Form of the soul, whose presence would indicate that the arguments are directed toward soul *qua* soul. Where he might naturally be expected to do so, in passages in which he lists Forms he and the interlocutors accept, or in the Final Argument where he actually discusses Forms in which souls participate (105b-107b), he fails to postulate such a Form of soul.

It makes more sense, then, to interpret (1) as (1b): the soul is the proportion as it is instantiated in each case by elements combined to form one particular being. Although it would be more usual to characterize *sounds* of a *harmonia* as beautiful and divine, if Simmias refers to the beauty of the perfect ratio itself, alternative (1b) looks perfectly plausible.<sup>15</sup> This alternative specifies that the *harmonia* is not the universal but the particular instantiation of that universal in each case, something along the lines of an immanent form. Such a reading of the thesis is familiar in the context of the dialogue: in the Final Argument for immortality of the soul (102a-107b) Socrates distinguishes the form in us ( $\tauὸ\ \dot{\epsilon}\nu\ \eta\mu\mu\tau\omega$ ) from the form in nature ( $\tauὸ\ \dot{\epsilon}\nu\ \tau\eta\ \varphi\mu\sigma\epsilon\iota$ , 103b5), a distinction often taken by commentators to be a distinction between immanent and transcendent Forms.<sup>16</sup>

If we adopt reading (1b) of the thesis, however, we adopt a weaker reading which is susceptible to one of Socrates' counterarguments, the one commonly labeled B.<sup>17</sup> Each soul is constituted by one attunement which is responsible for both the soul's being a soul and also its being morally *good or bad*, so that all souls ought to possess equal moral virtue. Yet we know from experience, as Socrates points out,

that persons vary widely in their goodness or wickedness. Socrates cannot seriously be worried about any version of the *harmonia* thesis that would entail such an implausible assumption.

In the end, then, neither version of interpretation (1), the view that the soul is the proportion itself—whether as a universal (1a) or as the individual instantiations of that universal ratio (1b)—produces a suitable nonmaterialist view that Simmias might have advanced. As a result, both interpretations (1) and (2) must be rejected.

A review of Aristotle's critical discussion in *De Anima* I.4 suggests another plausible solution. At 408a5-21 he raises objections to several versions of the *harmonia* thesis, including the version that I will defend here.<sup>18</sup> Aristotle first distinguishes two senses of 'harmonia': "if we speak of harmony by focusing on its two <main applications>, one, the most proper, pertains to magnitudes which have motion and position, <where that is> a compounding of them . . . and then too <there is harmony> as a proportion of the things mixed" (εἴτι δ' εἰ λέγομεν τὴν ἀρμονίαν εἰς δύο ἀποβλέποντες, κυριώτατα μέν, τῶν μεγεθῶν ἐν τοῖς ἔχουσι κίνησιν καὶ θέσιν, τὴν σύνθεσιν αὐτῶν . . . καὶ τὸν τῶν μεμιγμένων λόγον—408a5-9). The first alternative is clearly a straightforward materialist thesis, and as such is not relevant to our passage in the *Phaedo*. The second application, however, could be interpreted as either materialist or nonmaterialist. Going on, at 408a20-21 Aristotle raises a new question concerning the soul: is the soul identical with a ratio or is it something else that originates in the parts (πότερον οὖν ὁ λόγος ἔστιν ἡ ψυχὴ, ἢ μᾶλλον ἔτερόν τι οὖσα ἐγγίνεται τοῖς μέρεσιν)? Now we have a third alternative, other than the bodily parts themselves and the proportion itself in which these parts are assembled: the soul may be a distinct entity, nonidentical to the material body, which nevertheless has its origin in that body. Aristotle may be suggesting here that the *harmonia* is neither a material entity nor a causally independent immaterial substance along the lines of the Platonic soul, but instead some third kind of thing.<sup>19</sup> It has been argued in the literature that Aristotle himself holds this kind of conception of the soul, where the soul as form is an immaterial substance causally dependent upon the material body.<sup>20</sup> Hence, Plato could easily have been aware of, considered, and worried about a view of exactly this kind; and it is precisely the alternative to substance dualism that ought to worry him.

Thus, the best interpretation of the text is alternative (3): that the soul is an entity produced by the combination of elements but distinct from both the proportion and the elements themselves. This interpretation, however, does not settle a further question: while the

soul is produced by material elements, is it a material or nonmaterial entity? If it is nonmaterial, then the attunement model will be sufficiently relevantly similar to the Platonic substance dualist conception to challenge it: the *harmonia* possesses most of the properties attributed to the soul in the Argument from Affinity yet plainly fails to be immortal.

As I have shown above, an analysis of the text reveals that the relevant conception must be a nonmaterial one, with the result that Simmias' thesis is necessarily dualist. We have of course rejected a Platonic substance dualist interpretation, since this is exactly what Simmias' objection is intended to refute; still, there are two types of dualism that might apply. The first is epiphenomenalism: Taylor, for example, thinks that the soul is "given out" by the body as a kind of epiphenomenal by-product, an emergent entity identical neither to the body nor to the ratio of the elements.<sup>21</sup> This account, however, is not consistent with Simmias' fundamental assumptions about the active nature of the soul, since epiphenomenal mental properties lack any causal efficacy. According to the traditional theory of epiphenomenalism, mental properties, while they are caused by physical properties, are themselves causally inert. Only physical properties have effects.<sup>22</sup> Yet the participants in the conversation assume that the soul will have a number of distinct effects upon the body. Most notably, several times in the dialogue the soul is said to rule the body (80a7, 94b4-11, 94c9-d5); this relation between them is not only asserted in the Argument from Affinity but is actually the basis of one of Socrates' arguments against Simmias' thesis. Any version of the thesis would therefore be called upon to explain activities and effects of the soul attributed to it. Plainly, the epiphenomenalist interpretation of the thesis is inconsistent with both facts about attunements of musical instruments themselves, since the tuning of strings of such instruments has the effect of causing sympathetic vibration of the other strings under certain conditions, and also with the depiction of the soul in the *Phaedo*.

More consistent with both the text and facts about *harmoniai* is a second type of dualist interpretation, the subtype of supervenience called supervenient dualism or substance supervenience.<sup>23</sup> It will be useful to stop here and sketch briefly a general account of supervenience. The core idea is that entities are subject to covariation or "dependent-variation" of two sets of respects.<sup>24</sup> Put simply by McLaughlin,

the core idea of supervenience . . . is the idea of dependent-variation, where the dependency is of a purely modal sort. Variation in *A*-respects

depends on variation in *B*-respects in that *A*-respects *cannot* vary without variation in *B*-respects. . . . Supervenience, so understood, is a purely modal dependent-variation relation that is reflexive, transitive, and and neither symmetrical nor asymmetrical.<sup>25</sup>

The virtue of this definition is that it is not overly restrictive: it does not attempt to specify the causal relation of the sorts of respects involved, and it leaves open the nature of the *A*-respects which vary. Differences in *A*-respects might include, for example, differences in what types of *A*-properties exist, what *A*-events occur, and most useful for our purposes here, what *A*-types of entities exist. In other words, the relata of supervenient relations may possibly be members of different ontological categories.

Most current literature on supervenience and the mind is cast in terms of supervening properties, however, or at most supervening mental events, rather than the supervening substance so clearly indicated in the *Phaedo*.<sup>26</sup> We will need to recognize that Simmias is concerned not with properties alone but with the entities that are the possessors of those properties. His thesis is an extension of straightforward supervenient (property) dualism that asserts the existence of a discrete supervening substance which possesses certain properties. This is the case because, in his argument, Simmias attributes to the soul a set of properties that is contradictory to the properties of the body. The properties in question are of the sort that cannot belong simultaneously to one entity; unlike color predicates, for instance, materiality and nonmateriality, compositeness and noncompositeness, are not attributable to one and the same thing at any time *t* in the same respect. By Leibniz's Law, soul and body are distinct entities. This is precisely what substance supervenience is: it is the view that the soul is an immaterial substance which supervenes on the body, which is also a substance.<sup>27</sup> In the form of dualism I suggest here, one substance supervenes upon another if and only if the properties of one substance supervene upon the properties of the other substance. One property supervenes upon another only if the following causal nomological relation, formulated crisply by Christopher Shields, is true: "A group of properties  $\Psi$  strongly supervenes on a group of properties  $\Phi$  if and only if: (i) necessarily for any property  $F$  in  $\Psi$ , if  $x$  has  $F$ , then there exists a property  $G$  in  $\Phi$  such that  $x$  has  $G$ , and (ii) necessarily if any  $y$  has  $G$  it has  $F$ ."<sup>28</sup> If we extend this definition to entities that are subjects of the given properties, then any supervening substance will meet the general requirements for supervenience if (a) it is dependent upon the base substance and (b) its properties covary with those of the base substance.<sup>29</sup> This view differs

from supervenience in its most common contemporary form, which is a materialist view, in postulating the existence of the supervening substance itself, rather than one substance with two sets of properties. Because this substance's properties do depend upon, and covary with, those of the base substance, it is supervenient upon that base substance. The simultaneous existence of properties G and F, which appear to be properties of one substance, is sufficient to indicate the presence of two substances.

The *harmonia* as defined by Simmias meets both of requirements (i) and (ii) in the definition of supervenience above. This interpretation of Simmias' thesis provides a model of the soul that is distinct from both the material elements of the body and the proportion in which they are combined, yet is dependent for its existence upon the subvening material base and hence is mortal—the precise threat to the Platonic substantial soul that will be most difficult for him to refute.

That Simmias should present a supervenience conception of the *harmonia* is very reasonable given the views on *harmoniai* held by Philolaus, with whom both Simmias and Cebes have associated (61d6-7, 67). In Fragment 6, Philolaus discusses the ordering of limited and unlimited which, he says, would have been impossible “if a harmony had not come upon them, in whatever way it came to be” (*εἰ μὴ ἀρμονία ἐπεγένετο ὥτεινιῶν ἀν τρόπῳ ἐγένετο*, 8-9).<sup>30</sup> Philolaus is asserting here that at least certain *harmoniai* come upon or supervene upon elements that are ordered. In considering whether or not the soul is a *harmonia*, then, Philolaus would be inquiring about it as the kind of attunement that comes upon, or supervenes upon, the bodily elements.<sup>31</sup> Given his medical interests, he would most likely accept a materialist conception of the soul as *harmonia*.<sup>32</sup> In contrast, while Simmias turns Philolaus' view of *harmoniai* to his own purposes, the thesis in his presentation of it is patently dualist—as indeed it must be as a response to Socrates' argument for immortality. It is this conception of the soul as a substantial supervenient entity that is his argument's most troublesome counterexample.

The substance supervenience view I have presented above is a greater threat to Socrates' three arguments against Simmias' thesis than has usually been recognized. Two of them, more extensive and detailed than the brief first attack on the thesis, occur in a pronounced chiastic structure at 92e4-95a3.<sup>33</sup> Of these the first, A, argues that the soul cannot be a *harmonia* because the soul has causal powers that a *harmonia* lacks; the second, B, asserts on empirical grounds that facts about differing moral states of various souls are inconsistent with the soul's being a *harmonia*. Let us examine argument

B to show exactly how the substance supervenience interpretation of the thesis explains B's failure to refute Simmias.

#### 4. A Troublesome Counterargument to the Thesis

Of the arguments Socrates offers against the thesis, the last one, Argument B (93a11-c10), is the most extensive and complicated. The argument is a *reductio ad absurdum*, as the following reconstruction shows:

- B1 The soul is an attunement. (86b8-c1)
- B2 Every attunement is an attunement just as it's been tuned. (93a11-12)
- B3 An attunement that has been tuned to a greater or lesser degree (if this is possible) is a greater or lesser attunement. (93a14-b2)
- B4 No soul is more or less a soul than any other. (93b4-6)
- B5 Souls participate to varying degrees in goodness and badness, where goodness is attunement and badness is non-attunement. (93b8-c1, 93e8-9)
- B6 No attunement is more or less an attunement than any other (from B1 and B4; invalid). (93d1-4)
- B7 What is neither more nor less an attunement has been neither more nor less tuned. (93d6-7)
- B8 That which has been neither more nor less tuned participates in attunement to an equal degree. (93d9-10)
- B9 No soul has been more or less attuned than another. (93d12-e2)
- B10 Therefore, no soul participates more in attunement or nonattunement than any other. (93e4-5)
- B11 Therefore, no soul participates more than any other in goodness or badness. (93e7-10)
- B12 Therefore, the soul cannot be an attunement. (B11 contradicts B5)

Socrates is undecided in B3 about the possibility of degrees of attunement.<sup>34</sup> The clause at 93b1, *εἴπερ ἐνδέχεται τοῦτο γίγνεσθαι* (“if this is possible”), expresses doubts about the existence of degrees of attunement, but does not necessarily imply that the claim that follows—namely, that an attunement within an attunement is possi-

ble—is false. One commentator has contended that the premise is nonsensical, as it concerns degrees of attunement of an *attunement*, not of an instrument.<sup>35</sup> But it is more reasonable to accept the premise because Socrates is simply preparing ground for the question of virtue and vice in the soul, which is precisely the question of the possibility of further attunements within such attunements as already exist. This interior (or supervenient) attunement would be normative, while the outer (or subvenient) one would serve as the formal cause of the entity's being a soul.

The conclusion B6, that no attunement is more an attunement than any other, is the crux of the entire argument. For if the only definition of a *harmonia* that Plato accepts is that of a perfect proportion, then all souls as *harmoniai* will be perfect. On the other hand, if an attunement admits of degrees, then the soul may be imperfect and argument B fails in its attack on the attunement thesis, since there is no longer a *reductio*. But this conception does not allow for either vice or virtue in the soul, a result which Socrates finds untenable in light of empirical evidence of human nature to which he refers. Therefore Socrates rejects premise B11, as it contradicts the obvious fact that there are both good and bad souls. On the other hand, if there are degrees of attunement, perhaps vice and virtue may be accounted for differently. Socrates suggests at 93b8-c9 that virtue would be a further attunement within the first-order tuning, and that vice would be a lack of attunement altogether. Two scholars whose views I discuss below agree that two levels or kinds of attunement are indeed the solution to the question of better and worse souls, but they also point out that this solution renders the argument fallacious.

The real problem with the argument, according to both Taylor<sup>36</sup> and Gallop,<sup>37</sup> is that it contains a fallacy caused by the uses of the term *harmonia*, though they diagnose the fallacy differently. The former argues that the argument makes use of incomplete predicates: here the predicate “attuned” is left without a complement. The ambiguity arises because the text of the argument does not state how, or with respect to what elements, the attunement is performed. At 93a11-12, Socrates asks, “isn’t it natural for every attunement to be an attunement just as it’s been attuned?” (οὐχ οὕτως ἀρμονία πέφυκεν εἶναι ἐκάστη ἀρμονία ὡς ἂν ἀρμοσθῇ;). The implication is that any attunement possesses an inherent nature or character. Taylor’s own reconstruction of the argument incorporates this assumption as his premise B1: “Any ordering that is ordered *F*-ly is an *F* ordering.”<sup>38</sup> He therefore argues that the term ‘ordering’ (equivalent to ‘is attuned’) is an incomplete predicate which is correctly

completed by specifying the elements in it as orderings of some specific type of element. Once this incompleteness is recognized, the fallacy is apparent: there are two types of orderings, physical and psychic, which are implied in the argument. Souls are orderings of *physical* elements, but virtue and vice are orderings of *psychic* elements. It is possible, Taylor continues, for an entity to be organized in one of these respects but to lack order in the other respect at the same time. Plato has failed to recognize that different premises of his argument refer to two types of attunements, physical and psychic; and this equivocation causes the argument to be fallacious.

Gallop argues along similar lines that Socrates’ argument fails, but he attributes the problem to a different equivocation on ‘attunement.’ There are two types of *harmoniai* according to Gallop: there is attunement<sub>1</sub>, which is just the state of being tuned, and attunement<sub>2</sub>, or the state of being perfectly tuned. For example, an attunement in the first sense is simply a correct set of pitches, while the second sense of attunement requires that the pitches be not only the correct ones, but perfect in intonation. Every attunement<sub>1</sub> is as much an attunement as any other, but not every attunement<sub>1</sub> is also an attunement<sub>2</sub>. The result is that an attunement in the first sense might also participate in nonattunement by lacking an attunement in the second, more stringent sense. This model explicates vice and virtue easily: a good soul contains both the attunement<sub>1</sub> and the further attunement<sub>2</sub> within it; the bad soul, while possessing the first kind of attunement, lacks the second. Every soul is completely a soul in being attuned in the first sense, yet it may still have degrees of attunement in the second sense.

Socrates could reply in the following way to this line of reasoning: he might argue, as the Final Argument asserts (102d5-103c8), that opposites cannot admit each other. No soul could therefore admit both attunement and nonattunement. But as Gallop has already pointed out, attunement<sub>1</sub> and attunement<sub>2</sub> are not identical, so that the opposite of attunement<sub>2</sub> is not the opposite of attunement<sub>1</sub>. A soul could possess order in the first sense and disorder in the second without admitting opposites. Each soul, remaining an attunement, possesses a second attunement that determines its moral state.

While both Taylor and Gallop are correct that the argument contains an equivocation, neither remarks on the true fatal weakness of Argument B. Socrates misdirects his attack on Simmias’ thesis, assuming that the only kind of attunement he need consider is a material one; he refutes only interpretation (2) of the thesis, that the soul is the combination of elements in the correct proportion. It is indeed nonsensical to assume that one and the same attunement causes a mass of matter both to be a soul, and also to be a morally good soul,

since obvious facts about the differing moral states of existing souls are readily available to Socrates to disprove this view. The supervenient soul, however, is not material, and is not limited to the proportion of the material elements of the body. Possessing its own set of properties, it explains the common variation in moral character Socrates notes, but at the same time it is plainly mortal. The counterargument fails against this interpretation of the *harmonia* thesis, reopening the question of the soundness of the Argument from Affinity.

## 5. What Is at Stake: Why Plato Must Refute the Thesis

I asked at the opening of this essay why Socrates should have thought the thesis sufficiently interesting and threatening to warrant formulating three quite ingenious arguments against it. One might ask further why Simmias' thesis is paired with Cebes' analogy of the weaver and cloak at that point in the dialogue, immediately following the Argument from Affinity. With this question we open the discussion of certain broader Platonic concerns with causation.

The salient feature of both Simmias' thesis and Cebes' analogy is that they are causal analogies. Recall that Cebes suggests that the soul is to the body like a weaver to the cloak he produces: a weaver may make a number of cloaks, each individually weaker than the man himself and each in turn worn out with constant use. Yet at the end of his life, the weaver may die leaving behind a last cloak, not yet worn threadbare. Comparably, perhaps the soul, while stronger than the body it inhabits, may wear out a series of bodies yet die before or with the last body it inhabits. While the soul is ontologically independent of the body, and the body is causally dependent upon the soul, still the soul may not survive the death of the vehicle it occupies and thereby maintains as alive.

What is evident from the paired objections is that Plato is concerned at this point in the *Phaedo* not only with the nature of the soul, but also with problems of its causal origin. The reason for his concern is clear when one investigates the concept of causation Plato himself seems to accept in the *Phaedo*, a question recently discussed with admirable lucidity by David Sedley.<sup>39</sup> Very briefly, Sedley notes that the causal language of the *Phaedo* regularly employs the following three locutions: *αἰτία/αἰτίαι* (*aitia/aitiai*); *διά* (*dia*) with accusative, or causal dative; and *ποιεῖν* (*poiein*) as to cause or to make (F).<sup>40</sup> Taken together the expressions represent a unitary conception

of "cause," one of causes as *things* rather than as events, states of affairs, forces, etc. Plato holds a conception of a cause as a *thing responsible* for any given effect. It is his motivation for holding this conception that is relevant to my argument here: as Sedley argues, causes for Plato are best explained as functioning by causing their like. Evidence adduced for this claim lies in the fact that Plato on the one hand consistently rejects any principle based upon the assumption that one of a pair of opposites (F) could cause the other (not-F), and on the other hand maintains that what causes F must itself be F.<sup>41</sup> Plato's explanations of causation are thus not epistemological matters, but ontological ones. Exploiting legal analogies, Sedley proposes that we conceive Plato's causal theory as a transmission theory in which, for instance, what is hot causes other things to be hot by transmitting its heat, just as the cause of a murderous act is most clearly traced to murderous intent on the part of the perpetrator.

Now we are in a position to understand why Plato should have included both Simmias' and Cebes' causal analogies as objections to the Argument from Affinity, and especially why he should have given himself ample opportunity in the text for a vigorous refutation of the *harmonia* thesis. Because the thesis assumes that it is possible for a material entity, namely the lyre or the body, to cause an immaterial entity, the attunement or the soul, this thesis is an implacable enemy of Plato's entire causal theory. To allow the thesis to stand unrefuted would be to deny the central facts of causation on which he bases his arguments. An additional difficulty would be that the theory of Forms itself depends upon Plato's causal theory to explain the mechanism of sensibles' participation in the Forms. It is evident that this dual threat, both to the immaterial, immortal soul and to the causal mechanism Plato accepts, must be eliminated.

## 6. Conclusion

The thesis presented by Simmias has usually been dismissed by commentators as an easy target for Socrates; yet on the supervenient dualist interpretation I have proposed it presents a stronger challenge than had previously been thought. My interpretation makes explicit a dualist threat to Platonic substance dualism, as it is presented in the Affinity Argument, which Plato was clearly obliged to counter. The soul may have all of the properties attributed to it in the Affinity Argument yet fail to be immortal.<sup>42</sup>

## Notes

1. Here it should be noted that my remarks are limited to English-language commentators. Among these, a number disagree on whether or not the counterarguments are successful. C. C. W. Taylor, "The Argument in the *Phaedo* Concerning the Thesis That the Soul Is a *Harmonia*" (hereafter *Harmonia*), in *Essays in Ancient Greek Philosophy* II, ed. John P. Anton and Anthony Preus (Albany: State University of New York Press, 1983), 230, and David Gallop, *Plato: Phaedo* (Oxford: Clarendon, 1975), 165-67 point out that there is a fallacy in Argument B, though Taylor considers the short initial counterargument cogent against all forms of the thesis. On the other hand, W. F. Hicken, "Phaedo 93a11-94b3," *Classical Quarterly* 4 (1954), 22 believes the arguments do refute the thesis.

2. I am not the first to argue for a supervenience interpretation of the thesis: see Victor Caston's excellent and comprehensive treatment of epiphenomenalism in "Epiphenomenalisms, Ancient and Modern," *The Philosophical Review* 106/3 (1997): 309-63. My claims differ from his primarily in that (1) I do not consider the supervenience of the *harmonia* thesis to be epiphenomenalist; and (2) I take Simmias to be presenting a *dualist* thesis, since he is presenting a counterexample analogous to a dualist view of the soul in the Affinity Argument.

3. There are two possibilities for the extent of the argument proper: it may be complete at the conclusion of 80b1-10, or it may end at 84b5-8. I consider the necessary argument establishing immortality to be complete at 80b, while the section from 80b10-84b4 elaborates and mythologizes about the various fates of the different sorts of souls.

4. All translations, with minor modifications, are taken from Gallop's commentary. All Greek is that of Burnet, *Plato's Phaedo* (Oxford: Clarendon, [1911] 1998).

5. While the argument is fundamentally analogical, certain members of the cluster of properties attributed to the first kind of being, the Forms, are logically prior to others, so that the argument is not strictly so. Only one property—constancy—is directly attributed to the soul in the argument. The property of noncompositeness is deduced from constancy, since the lack of parts is a necessary condition for constancy, rather than constancy a condition of being noncomposite. Similarly, noncompositeness is logically prior to indissolubility, which follows because there is no means of breaking apart what does not have components. For the most recent and thorough treatment of the argument, see David Apolloni, "Plato's Affinity Argument for the Immortality of the Soul," *Journal of the History of Philosophy* 34/1 (1996): 5-32.

6. Simmias' use of *ἀσώματον* (*asômaton*) is, according to LSJ, one of its first appearances; it may indeed be one of Plato's neologisms. Whether lacking material body implies any other of the properties we might now assume accompany this property (lack of mass, etc.) is impossible to determine from Simmias' presentation.

7. While there is no explicit attribution of compositeness to the *harmonia* here, it is formally acknowledged at 92a8 and 93a1 in Socrates' counterarguments.

8. It is perhaps surprising that Socrates does not seize upon this fact and reject Simmias' argument out of hand as irrelevant, on grounds that the *harmonia* is not a legitimate counterexample. Plainly, Socrates wishes to pursue this refutation.

9. See the note on this passage in C. J. Rowe, ed., *Plato: Phaedo* (Cambridge: Cambridge University Press, 1993). Rowe points out that "it is the capacity to produce concordant music which would naturally be said both to be 'in' the tuned lyre, and to be *πάγκαλόν τι καὶ θεῖον*" (203). I take it this dative is best characterized as locative; on the use of *ἐν* with dative see LSJ and also Smyth §1687, as well as the note on 85e3 in Burnet's commentary.

10. Those who hold the thesis to be nonmaterialist are R. S. Black, *Plato's Phaedo* (London: Routledge & Kegan-Paul, 1955); R. Hackforth, *Plato's Phaedo* (Cambridge: Cambridge University Press, 1972); Theodore Scaltsas, "Soul as Attunement: An Analogy or A Model?", *Greek Studies in the Philosophy and History of Science*, ed. Panetelis Nicolacopoulos (Dordrecht: Kluwer, 1990); and Taylor, *Harmonia*. Gallop, *Plato's Phaedo*, and R. D. Archer-Hind, *The Phaedo of Plato* (New York: Arno Press, 1973), do not commit themselves concerning its materialist or nonmaterialist nature.

11. Taylor, *Harmonia*, 218.

12. This is inconsistent with the obvious truth that a *harmonia* in the sense of an attunement of a specific lyre is audible when the lyre is being played. Yet Simmias clearly does not think that its possible audibility is inconsistent with its immateriality and divinity.

13. So Hackforth, *Plato's Phaedo*, 102, and Gallop, *Plato: Phaedo*, 148. The latter finds the suggestion that "we" refers to people in general more agreeable than the possibility that it refers to some sort of heterodox Pythagorean view. Also, Carl A. Huffman in *Philolaus of Croton: Pythagorean and Presocratic* (Cambridge: Cambridge University Press, 1993), 73-74 and 328-32 discusses Philolaus' view of *harmonia* and number. Since Simmias and Cebes are said to have been with Philolaus (61d6-7), which Cebes admits (61e7), these views are most likely relevant to the discussion of the soul as *harmonia* in the *Phaedo*.

14. See Taylor's discussion of this alternative, *Harmonia*, 223.

15. If Simmias does indeed refer to the sounds of the attunement as the lyre is plucked, his concomitant characterization of the *harmonia* as incorporeal would be puzzling, since the sounds are audible and hence accessible by at least one of the senses, which presumably any entity lacking body completely could not be. It was on the grounds that soul is invisible that it was assimilated to the Forms in the Affinity Argument, though Socrates quite clearly is trying to establish not the soul's invisibility but its incorporeality. Simmias himself makes the ontological status of the *harmonia* explicit when he states that it is *ἀσώματον* (incorporeal).

16. Immanent forms in the *Phaedo* have been a bone of contention among Plato scholars. For commentators who accept immanent forms, see David Keyt, "The Fallacies in *Phaedo* 102a-107b," *Phronesis* 8/2 (1963), 168 nn. 1 and 2, 169 n. 1; Gregory Vlastos, "Reasons and Causes in the *Phaedo*," *Philosophical Review* 78 (1969), 318 n. 70; and R. Hackforth, *Plato's Phaedo*, 149 n. 4 and 150 n. 1. On the other hand, W. J. Verdenius, "Notes on Plato's *Phaedo*," *Mnemosyne* 11 (1957), 232, and D. O'Brien, "The Last Argument of Plato's *Phaedo*," *Classical Quarterly* 17 (1967), 204, 208, do not. To my mind the most sensible assessment of the texts is Gallop's: he remarks in his commentary that "no consistent distinction between 'immanent' and 'transcendent' Forms can be founded upon Plato's terminology," *Plato: Phaedo*, 195. The question arises because of its relevance to reading (1b); the *harmonia* thesis must have been worrisome to Plato in this version because it so closely resembles his own characterization of the differences between, for example, largeness in us and Largeness itself. Commentators have generally thought that the suggestions that largeness in us might perish upon the advance of the small (102e2) and that any of the opposites might perish (103a2) only make sense if Socrates refers to immanent forms.

17. See below, Section 5.

18. Taylor, *Harmonia*, 221, Gallop, *Plato: Phaedo*, 148.

19. See the excellent discussion in T. H. Irwin, *Aristotle's First Principles* (Oxford: Oxford University Press, 1988), 281.

20. Christopher Shields argues for this view in "Soul and Body in Aristotle," *Oxford Studies in Ancient Philosophy* 6 (1988), 103-37. As he notes, his own view counters that of many commentators, who are inclined to argue that Aristotle is a weak materialist (105 n. 3).

21. See A. E. Taylor, *Plato: The Man and His Work* (New York: Meridian Books, 1956), 194. H. B. Gottschalk, "Soul as *Harmonia*," *Phronesis* 16 (1971), 181 criticizes this depiction. He notes the same worrisome ambiguity I do in the dependence of soul upon the body despite properties it possesses that seem inconsistent with its being a dependent entity (the "positivism and mysticism" of Plato are difficult to reconcile). Gottschalk explains this combination as a result of Plato's adding, to the assumption of the soul's divinity in the Argument from Affinity, the "positivism" implicit in making the soul dependent on the body (195).

22. An excellent brief introduction to epiphenomenalist views is the entry on epiphenomenalism by Brian P. McLaughlin in *A Companion to the Philosophy of Mind*, ed. Samuel Guttenplan (Oxford: Blackwell, 1995), 277-88.

23. Much of the seminal literature on supervenience is the work of Jaegwon Kim. He has discussed both epiphenomenalism (briefly) and supervenience in *Philosophy of Mind* (Boulder: Westview, 1996), 51 f. See also his "Postscript on Supervenience," 165 n. 5 in *Supervenience and Mind* (Cambridge: Cambridge University Press, 1993). For an account of the kind of supervenient dualism I defend here, see Kim's "Causality, Identity, and Supervenience in the Mind-Body Problem" in *Midwest Studies in Philosophy* IV, ed. Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein (Minneapolis: Uni-

versity of Minnesota Press, 1979), 31-49. Although Kim later claims that supervenience is "fundamentally" a materialist view, he thinks the question of reducibility, which would settle the issue of materiality or nonmateriality of the supervening mental properties, is so difficult that it should not even be a component of the definition of supervenience. If we hold that properties are themselves nonmaterial entities (are abstract, etc.) then we have a plausible basis for defining supervenience as a nonmaterialist view. This is approximately the interpretation of Simmias' thesis that I find plausible and think Plato himself worried about.

24. Kim commonly defines supervenience in terms of co-variation of, for example, properties. In contrast, Brian McLaughlin, "Varieties of Supervenience" in *Supervenience: New Essays*, ed. Elias E. Savellos and Ümit D. Yalçın (Cambridge: Cambridge University Press, 1995), 18, prefers the term 'dependent-variation' over 'co-variation' though the terms have (he claims) the same meaning, because 'dependent-variation' does not give the false impression that the supervenience relation implies either symmetry or asymmetry of variation. Complaining of the lack of a general definition of supervenience in the literature, rather than the definitions of its numerous subtypes that are the focus of most of the argumentation, McLaughlin offers a general account of purely modal dependent-variation in two respects on which I draw here.

25. McLaughlin, "Varieties of Supervenience," 18. Notice that, unlike epiphenomenalism, supervenience does not necessarily include assumptions about the direction of causation between the subvenient and supervenient entities.

26. See Kim's recent work, including "The Mind-Body Problem: Taking Stock after Forty Years," *Philosophical Perspectives* 11: *Mind, Causation, and World* (1997): 185-207.

27. The following draws upon Shields' account of Aristotle's supervenient dualism in "Soul and Body in Aristotle" because Simmias' thesis depends in a similar way upon the existence of two discrete substances (Shields' ground for attributing substantiality to the soul in Aristotle, and hence the dualism of Shields' view, is that the soul as form is substantial).

28. This definition is Shields' in "Soul and Body in Aristotle," 133.

29. These requirements are discussed at length by Kim in "Supervenience as a Philosophical Concept," *Metaphilosophy* 21 (1990): 1-27. Also, see Caston's thorough, detailed treatment of epiphenomenalism in "Epiphenomenalism, Ancient and Modern," especially 319-25. Caston explains that the *harmonia* metaphor "essentially expresses a supervenience thesis" (322) but does not consider the question of the *harmonia* as substance, and thus does not discuss supervenience in this case as dualist, rather than covariant. In fact, he states that supervenience "is simply a form of (nonsymmetric) covariation" (313); certainly this is true of the contemporary form of the theory, but there is evidence in Simmias' presentation of the thesis that he takes it necessarily to be dualist. Hence, what Socrates considers is subject to constraints no modern theory of supervenience need take into account.

30. See Huffman, 123, for the Greek text and translation. Huffman discusses the line at length (138-41), treating the verb *ἐπεγένετο* (*epegeneto*) in special detail (140-41). He claims that “in his use of this word, Philolaus is clearly emphasizing that *harmonia* is something that ‘comes upon’ or ‘supervenes on’ the two beginnings that he has just postulated, limiters and unlimiteds” (140).

31. Huffman points out that the soul, as the attunement that orders the limiters and unlimiteds in the body, would “be identified with the *harmonia* that ‘supervenes’ on limiters and unlimiteds as described in F6” (328). He also notes that F1 and F6 seem to imply that attunements “are always attunements of something and not to be regarded as independent entities, as does the vague description of a *harmonia* supervening ‘in whatever way it does’ in F6” (328).

32. Indeed, as Huffman concludes after a meticulous survey of both external evidence and the attested fragments themselves (III.6, 307-32), “it appears very likely that Philolaus thought of the soul in largely material terms as a group of constantly moving elements in attunement located in the heart” (329).

33. I follow Gallop, *Plato: Phaedo*, 158 (after Hicken) here. The structure of the argument is A-B-B-A: premises for Argument A (92e4-93a10); premises for Argument B (93a11-c10); Argument B (93d1-94b3); Argument A (94b4-95a3).

34. Hicken, “*Phaedo* 93a11-94b3,” 17.

35. Gallop, *Plato: Phaedo*, 160.

36. Taylor, *Harmonia*, 227-30.

37. Gallop, *Plato: Phaedo*, 164 f.

38. Taylor, *Harmonia*, 227.

39. See David Sedley, “Platonic Causes,” *Phronesis* 43 (1998): 114-32.

40. Sedley, “Platonic Causes,” 115. He claims that these expressions are, as indeed they seem to be, used interchangeably throughout the dialogue.

41. Whence, of course, the difficulties of self-predication and the Forms and the resulting voluminous literature.

42. I am grateful to Samuel Kimball, Christopher Shields, Nicholas D. Smith, William Prior, and the anonymous referees for this journal for valuable, incisive comments on earlier drafts. The audience at the Eastern Division Meetings of the American Philosophical Association, December, 1998 also provided helpful criticisms, especially my commentator Hugh Benson and Dory Scaltsas. All errors are, of course, my own.

## II. The Tripartite Soul

## Plato's Theory of Human Motivation

John M. Cooper

### I

Everyone knows that in the *Republic* Plato advances the theory that the soul has three independent parts (reason, spirit, and appetite, as they are usually called in English). Using this theory he constructs an account of the human virtues: each of the three parts of the soul has its own special role to play in a human being's life, and virtue, for us, consists in each of them playing its own role fully and in harmony with the others. Thus human virtue taken as a whole, according to the *Republic*, is a complex interrelationship among three separate psychological elements, each of which has its own indispensable contribution to make.

Now this theory of virtue contrasts sharply with the Socratic theory found, for example, in the *Protagoras*.<sup>1</sup> According to the Socratic theory virtue is essentially a property of the intellect (and never mind what other parts of the soul there may be). That Plato in the *Republic* is self-consciously rejecting this Socratic theory is by now well accepted; and most philosophical readers no doubt agree that the *Republic*'s theory is a distinct improvement. Even if knowledge by itself does motivate action, as Socrates evidently though obscurely assumed, there are surely other motivating factors as well, and being virtuous must therefore partly consist in having these other factors, whatever they may be, in some special condition or other. After all, it will be agreed by all parties that to be

virtuous is to have one's practical attitudes and dispositions—whatever it is that affects one's actions and the ways one is inclined to act—structured in some special way; the virtuous person's practical attitudes must be such as always to produce the (or a) virtuous and right action in the given circumstances. And if not only one's *thoughts* about what is good and bad, but also ways one *feels* about things (whether or not those are also ways one thinks about them) constitute practical attitudes affecting the ways one is inclined to act, then obviously virtue must be something more complex than the Socratic theory represents it as being. It must involve not just well-informed, correct thought about what is good and what bad for a person, but also certain specific states of feeling about these matters as well. From this perspective Plato's *Republic* theory can be seen as a stage in the progression from Socratic rationalism to the Aristotelian theory that moral virtue is an interfusion of reason and desire—reason having the truth about the ends of life and how to achieve them, and desire embodying these truths so that the person habitually wants just the things that reason says are worth pursuing.

This picture, though I believe it correct as far as it goes, does push to one side the details of Plato's theory of what motivates human action; his view that there are *three* parts of the soul is treated as an uninteresting oddity, wisely omitted by Aristotle from his account of virtue.<sup>2</sup> Even Platonic scholars, who as a group are not noted for their sensitivity to Platonic error, sometimes admit to being embarrassed by this part of Plato's theory;<sup>3</sup> and it is indeed not easy to resist every clever freshman's impression that Plato held there were precisely *three* parts of the human soul only because he needed three in order to push through the argument launched at the beginning of the *Republic*'s second book. Assuming that justice in the state must be the same as justice in the individual, and having plausibly argued that justice in the state requires the recognition of three separate classes of citizen making three different contributions to the social welfare, Plato is committed to there being correspondingly three separate parts of the soul performing three different functions in the organization of the just individual's life. Does he then simply force the facts of human psychology to fit theoretical preconceptions derived from these other parts of his argument? Or does he after all produce cogent independent reasons, based in unbiased reflection on facts about individual human beings, for adopting this theory?

In this chapter I want to argue that when understood properly Plato's theory presents in a quite subtle and interesting way undoubted facts about the psychology of human motivation, and that this theory accounts

for some central features of human beings better than other later theories are able to do. Though there is no denying that Plato's way of parceling out the different forms of human motivation seems at first rather primitive, and is at all events somewhat alien to our way of thinking, it has a powerful rationale of its own that is worth exploring. In fact, there is good reason to think that for Plato, despite the order of exposition, the view that justice requires three distinct social classes rather derived support from than gave support to the theory that the soul has three independent parts.<sup>4</sup> It is the psychological theory that Plato thought more firmly anchored in the facts. If this is right then in reconstructing the argument of the *Republic* one must give the psychological theory pride of place.

## II

It is evident that the question "How many distinct parts has the soul?" can only have a clear sense and receive a definite, nonarbitrary answer if it is understood against the background of some well-defined theoretical interest. Plato makes clear enough his own point of view when he first raises his question about the parts of the soul (435b-c). He asks whether there are in each of us three things corresponding to the three kinds of person the recognition and proper use of which he has argued is essential to good order in a city. Now it is by what the three kinds of person do or don't do that the city's corporate life is determined—what it does and doesn't attempt to do, what its overall aims are, what it succeeds or fails in doing, whether for good or ill. Similarly, then, the question how many parts the soul has, and whether it has three parts, as the city does, is the question how many distinct types of psychological input go to determining a person's choices and voluntary actions, that is, the pattern of his life in general. Plato's theory that there are three parts is, roughly, the theory that there are three psychological determinants of choice and voluntary action.

Now there is a familiar modern theory, going back to Hobbes,<sup>5</sup> that a person's actions are the joint product of his (relevant) beliefs and desires and nothing else—desire providing the original motive force and belief factual information about how to act in order to satisfy desire. On this theory there are two sorts of determinants of action, belief and desire, one of which (desire) is the exclusive source of motivation while the other (belief) contributes only factual information, but no additional im-

pulse to action. There is a misleading superficial similarity between this theory and Plato's. For on Plato's theory (as indeed on Aristotle's) in some ways the basic division is between reason on the one side ( $\tau\ddot{o}\lambda\omega\gamma\iota\sigma\tau\iota\chi\delta\acute{o}\nu$ , literally the calculating part) and appetite and spirit together on the other. And since reason is assigned the job in the soul of being wise and knowing the truth (441e4-5, 442c5-8) it seems at first sight not unnatural to think of it as playing the same role as belief plays on this modern theory; Plato would then be admitting *one* source of information but (surprisingly) dividing motivating desires into two classes, the appetitive ones and those issuing from "spirit."

But this interpretation is incorrect. On Plato's theory all three of the parts, reason as well as appetite and spirit, are independent sources of motivation; the contrast between reason and the other two is not really akin to the modern theory's distinction between inert, purely factual belief and motivating desire. This fact does not emerge with perfect explicitness until the ninth book, where Socrates advances the claim that "as there are three parts, there are also three kinds of pleasure, one peculiar to each part, and so with desires" (580d7-8, tr. Grube). That is to say, there are desires of reason as well as bodily appetites and impulses of a spirited nature. Strikingly, the word for "desires" here, *ἐπιθυμίαι*, is the word used throughout the *Republic* as the generic name for the urgent bodily appetites (thirst, hunger and sexual desire) that serve as paradigms for the third part of the soul, *τὸ ἐπιθυμητικόν*, which is so named after them. The desires of reason are thus implied to be strong impulses of some kind which we experience simply and directly because we possess the power of reason, the power to figure things out (*λογίζεσθαι*) and know the truth. Socrates specifies one of these desires a little later in the ninth book when he says that "It is obvious to anyone that the part by which we learn is always wholly straining to know where the truth lies" (581b5-6, tr. Grube). On Socrates' view, then, merely in virtue of having mind—of having the capacity to inquire into and discover the truth—we possess the *desire* to do these things. According to Socrates, the desire to know the truth cannot be wholly explained as the outcome, say, of our discovery that knowing the truth helps us to advance the goals which our appetites, or other reason-independent desires, incline us toward; nor does it result simply from the discovery that, to use Kantian terminology, our sensibility is so constituted that we happen to find knowing the truth (or thinking we know it) gratifying. One's desire to know the truth might be strengthened in these ways, but there always remains an irreducible desire for knowledge that is not dependent on an interplay between rea-

son and other aspects of our nature. This desire is an original constituent of human nature, as much so as our appetites themselves, or our sensibility in general. Socrates admits that not every person feels it as strongly or as steadily as some people do, and that some people's actions are motivated by it more often than others' are, but it must be active to some extent in everyone's life; the consequence of supposing that someone never experienced this desire would be that that person had no mind at all, and so was not a human being after all.

But intellectual curiosity is not the only desire Socrates attributes to reason. For in the fourth book he assigns to reason a double job: to know the truth and to rule (*ἀρχεῖν*, 441e4, 442c5) in the light of it. For reason to rule here takes the form of its deciding on its own authority what is the best thing to do, issuing injunctions (442c6, *ταῦτα παρήγγελεν*), and seeing to it that the required action is undertaken. And just as Socrates makes the desire for knowledge—that is, the desire which leads reason to perform one part of its natural job—the direct consequence of our rational nature, so, I believe it can be shown, he also assigns to reason an inherent desire to perform the other part of its natural job, that of ruling.<sup>6</sup>

That according to Socrates human reason has, so to speak, an innate taste for ruling, just as it has an innate taste for knowing, can be most convincingly brought out by considering the way in which he attempts to argue the distinctness of reason from appetite. Notoriously, he thinks that the fact that sometimes reason opposes appetite shows that they must be distinct parts of the soul: his example is an incompletely described case where a man is thirsty, that is (as he says) desires, yearns for, and has an impulse to, drink ( $\beta\omega\acute{u}\lambda\epsilon\tau\alpha\iota\pi\iota\epsilon\iota\nu\kappa\alpha\iota\tau\omega\acute{u}\tau\alpha\iota\delta\acute{e}\gamma\epsilon\tau\alpha\iota\kappa\alpha\iota\epsilon\pi\iota\tau\omega\acute{u}\delta\omega\acute{m}\alpha$  439bl), but something else, which comes from reasoning ( $\acute{e}\chi\lambda\omega\acute{u}\iota\sigma\mu\o\iota\delta\iota\mu\alpha$  d1), pulls him back ( $\alpha\acute{u}\theta\acute{e}\lambda\kappa\epsilon\iota\delta\iota\mu\alpha$  b3) and forbids drinking ( $\kappa\omega\acute{u}\lambda\kappa\epsilon\iota\pi\iota\epsilon\iota\nu$ , cf. c6-7, 9). There are several unclarities about this example (and, indeed, the other cases of conflict that Socrates argues prove that there are distinct parts of the soul). The text is not explicit as to whether in saying that reason opposes appetite he means merely that the object which appetite goes for reason rejects, or rather that reason in rejecting the object also addresses and opposes the appetite itself. A moment's reflection shows that he had better have in mind the stronger thesis if he is to have any chance of ending up with precisely three parts (and not indefinitely many), and, as we shall see, it is quite clear that this is how he conceives the opposition of *spirit* and appetite (cf. 439e-440a).<sup>7</sup> So, following T. H. Irwin,<sup>8</sup> I shall interpret him as claiming that because reason sometimes rejects an appetite—i.e., insists that an app-.

tite is *not* to be acted on, that it does *not* constitute a reason, say, to drink whatever liquid may be in question—reason and appetite must be distinct. Still, however that may be, the text does make it clear that Socrates is conceiving reason as a *force* which works counter to appetite, pulling the agent back from what appetite pushes him toward. He draws an analogy (439b8-11) between what goes on inside this thirsty man and what happens when an archer draws his bow: just as the archer's one hand pulls the bow to him while the other hand pushes it away, so thirst moves him toward the drink while reason pulls him back from it. This shows that already in Book IV reason is being conceived as itself a source of desires, of motivating conditions. But clearly enough, the desire of reason at work in this case is not plausibly represented as what I've been calling intellectual curiosity; so, apparently, reason has other desires than the desire to know the truth. That the rational desire at work in the thirsty man's case is a form of the desire of reason to *rule*, will emerge from consideration of an objection that might be raised against Socrates's use here of the archer analogy.

It might be claimed that this analogy is misleading, and that whatever can legitimately be meant by saying in a case like this that reason pulls one back, does not justify the attribution to reason of any motivating force of its own, on all fours with that belonging to thirst. For suppose I am thirsty but know the only available water is boobytrapped so that I'll get a painful electric shock upon coming into contact with it. I thereupon hold back because I want to avoid this pain. Here, although it may be fair enough to say that reason restrains me, this does not imply that reason is the original source of any motivating desire not to drink; what motivates me to abstain from drinking is my aversion to pain. If one is to speak in terms of forces here at all, then, the forces that come into conflict are these two physical desires, thirst and the aversion to pain, neither of which has its origin in any capacity for reasoning. If this is an example of the sort of conflict Socrates has in mind, then he is not entitled to treat reason as a motivating force on its own, and so the question doesn't arise, what *kind* of desire of reason is working here.

I think, however, that in the case as just described, Socrates would actually agree that only the aversion to pain motivated abstention; if one takes desires simply as givens and limits oneself to working out how to satisfy them, allowing, for example, whichever is the stronger to determine one's action, or working out and following some scheme whereby the totality of one's desires may be satisfied as fully as possible, then there seems no doubt that one's calculations have not contributed any-

thing to the already existing motives to action. In a similar case (554d9-e1) Socrates speaks of opposed appetites, not a conflict of reason and appetite. Presumably, then, he thinks that not every case of conflict is like this, even if some may be. In fact there seems no doubt that on the *Republic*'s scheme reason is taken to be capable of deciding on its own theoretical grounds which ends are worth pursuing, and does not merely (as the calculating just described does) provide the means to, or work out some balance among, appetitively or otherwise given ends.<sup>9</sup> When it proposes an end on its own authority, Socrates evidently thinks, reason also, at least sometimes, contributes a *desire* of its own (the desire to achieve that end), and this is an additional motivating force, over and above whatever other kinds of desire may also be operating. Perhaps, therefore, the case Socrates has in mind is one where such a reason-generated desire comes into conflict with an appetite.

If so, the conflict Socrates has in mind is of the following sort. The thirsty man has worked out (or at any rate holds) on grounds of reason that health is a good thing, a more important good than the momentary gratification of appetite. He also desires to preserve his health *because* it is a good thing, and this desire (a product of reason) conflicts with his thirst; in the case envisaged the desire of reason wins out, and the man abstains. Now presumably Socrates does not hold that the desire for health is part of the original constitution of human nature (as the desire for knowledge is apparently thought to be); it is instead the consequence of a higher-order desire for good, as such, together with the grounds, whatever they are, on which health is thought to be a good. So what is inherent in reason is the desire for good, as such—not the desire for any particular good. And Plato does of course in the *Republic* (e.g., 605d11-e1), as elsewhere, speak of this desire as one which all human beings have. We are now, however, in a position to say something more illuminating about the status of this desire in human life than simply that everyone has it. The desire for the good can now be seen as equivalent to the desire on the part of reason to work out the ends of life on its own and to achieve them. Reason wants to do these things *on its own*, that is, without treating the fact that one has an attachment for a thing grounded in appetite or spirit or any other source of desire there may be as a ground for pursuing it. Thus the claim that the desire for the good is inherent in reason itself amounts to the claim that anyone who possesses the power of reason wants to think out on his own, on purely rational grounds, what goals to pursue in life, and to achieve those goals. He wants, in other words, reason to rule in his life.<sup>10</sup>

## III

It is as independent influences on action, sometimes in conflict, sometimes in harmony with desires of reason, that appetite and spirit figure in Socrates' theory. Let us consider appetite first. Unfortunately Plato is not careful to give a systematic general description of the sorts of desires that he counts as appetitive; in the fourth book he focuses simply on what he calls the "clearest" (*ἐναργεστάτας*, 437d3) instances of what he has in mind, thirst and hunger and (he adds a little later, 439d6) sexual desire, allowing his treatment of these examples to substitute for a general account. Indeed, he insists on quite a narrow construction of even these appetites: if I am thirsty for beer or hungry for chocolate these desires are not, he says, correctly classified merely as thirst or hunger (cf. 437d8-e6). They are thirst or hunger modified by some addition (e7-8). He does not indicate exactly what the relevant addition is, but presumably it is an acquired liking for the taste of beer and chocolate respectively (together perhaps, in the former case, with a liking for the way beer makes me feel). Thirst, just by itself, unmodified by these or other additions, is simply for drink, hunger simply for food.

These examples, and this treatment of them, might suggest that Plato limits the appetitive desires just to the basic recurrent biological urges, and indeed only to that part of them which is primitive and unmodified by the effects of experience. On the other hand even in this passage he refers several times to "other appetites" than these three (439d7, 436a11 f., 437d 2-3);<sup>11</sup> and he gives an interesting example of such an appetite when in telling the story of Leontius, he refers to Leontius' "appetite" (*ἐπιθυμία*, 440a1) to look at some corpses piled by the road. In the same context, arguing that spirit never allies itself with appetite (440b4-7), he points out that a decent man if he thinks he has been in the wrong cannot become angry even if he is subjected, in just retaliation, to hunger and cold and other such physical deprivations (c1-5)—so that being shivering cold is or gives rise to an appetite with which spirit refuses to ally itself. Later, especially in Books VIII and IX, the love of money is repeatedly treated as an appetite—indeed where we might expect the expression "appetitive part" Plato fairly often in these books writes "money-loving" (*φιλοχρήματον*) or "profit-loving" (*φιλοκερδές*) part instead.<sup>12</sup> And the democratic man, whose principle of life is said (561b2-c3) to be to give free and equal scope to each of his appetites, is credited not merely with a large variety of particular appetites for many different kinds of food and drink and sex, but also with appetites for

various athletic and political pursuits and even, on occasion, for, as he imagines it, doing a little philosophy (561c-d).<sup>13</sup> Thus desires that embody modifications of the basic appetites for sex, drink, and food (e.g., the desire for lobster) are nonetheless still appetitive desires; likewise physical desires which would not ordinarily be called appetites, such as the desire when cold to be warmed up, or the aversion to pain, count as appetitive, as do ghoulish impulses like Leontius' for looking at dead bodies. So also the love of money and the liking for physical exercise. *Some* kind of liking for such things as political activity and dabbling at philosophy also counts. What principle of unity is Plato relying on here? Is there really one at all?

In considering this question let us begin where Plato does, with the recurrent biological urges for food, drink, and sex. Socrates' first concern is to convince his interlocutors that there are *two* independent sources of motivation, reason on the one side and appetite on the other. For this purpose it is essential to choose examples that are undoubted cases of desires motivating action but where there is equally no doubt that these are desires not having reason as their source. Desires for specific kinds of food or drink or acquired likings of any kind will therefore not do: the generation of these desires obviously involves the use of *some* power to reason, at least to the extent of noticing and remembering the effects on oneself of various eaten and drunk objects or various external conditions and activities. And even if, as I believe Plato would ultimately want to argue, these rational powers ought not in such cases to be construed as belonging to the part of the soul which he calls "reason," to assume that at this point would certainly be confusing. Nor is it necessary for him to do so. There seems no doubt that hunger and thirst, understood as simple urges for food and drink, arise wholly from physiological causes (cf. *διὰ παθημάτων τε καὶ νοσημάτων παραγίγνεται*, 439d1-2), without any intervention from or detour through reason, not even through these equivocal processes of noticing and remembering. It is equally clear that these things have a direct motivating influence on action, as the fact of conflict to which Plato appeals very clearly indicates. Hence by concentrating on hunger and thirst as his "clearest cases" he can convincingly demonstrate the existence of motivating desires that work altogether independently of reasoning of whatever sort. And that is all he wishes, and needs, to show at this point in the argument.

Once it is established that there *is* such a source of motivating desires, independent of reason, it is not difficult to recognize other desires besides the recurrent biological urges as having essentially the same

status. Thus there are other desires besides hunger, thirst, and sexual appetite that are based on physical and physiological causes: e.g., the desire to be warmed up when cold, or, in general, the aversion to pain. And certain other more complex desires can be treated as transformations of these and other such appetites: thus all particular likes and dislikes in food and drink. Some tastes are simply found to be pleasant, and those which are, generate, by straightforward physical causation, desires for them. These desires, in turn, give rise, provided one knows what external objects need to be manipulated in order to get the pleasure, to further desires for those objects themselves. (This is what justifies Plato in counting the desire for money as appetitive: see 580e5 ff.) So, even though what we (but, as I have indicated, not Plato) would classify as rational powers may be involved in the constitution of such further desires it is not at all events *motivating* reason, but only calculation undertaken in the interest of the appetitive goal of physical gratification. In this way, beginning from simple hunger and thirst, we can explain why thirst for beer or hunger for chocolate should count as desires of the same basic type: all these desires rest ultimately on brute facts about our bodily constitution and about the means by which pleasurable bodily states may be caused.

It does not seem, however, that Plato means to limit the appetitive desires to those whose origin lies in such facts about our bodily constitution. At any rate, Leontius' ghoulish desire to look at dead bodies, or the democratic man's liking for philosophical dabbling might seem poor candidates for this kind of treatment. It might give one pause, however, that when Leontius is overcome by his appetite and his spirit intervenes to chastise him for overruling his reason and gaping at the bodies, Socrates says spirit places the blame on his eyes: "Go ahead and look," Leontius is quoted as saying, "you wretched things; get your fill of that lovely scene" (440a2-3). This might suggest that according to Plato it is the constitution of eyes (at any rate Leontius') that makes looking at corpses so fascinating to him: in the same way as my particular tastebuds are responsible for the fact that I enjoy the taste of orange juice, Leontius' eyes give him a pleasure caused by the sight of dead bodies. It would be difficult, however, to sustain this suggestion: it seems certain (unless we are to understand Leontius' attraction as straightforwardly sexual) that whatever it is about dead bodies that so interests him has something to do with some way in which he is thinking of them—some thrill-inducing contrast between living, animated human beings and these limp and broken figures, say—and it seems too much to believe that

anyone's eyes are naturally so constituted as to be given some pleasure by being exposed to dead bodies *when* so conceived. Leontius' imagination is at work here, and, surely it, rather than the eyes, is the most important source of the pleasure he is seeking.

Still, the workings of the imagination might be thought of as the source of pleasure in the same way as the bodily senses are. A person simply *finds* certain imaginings interesting or amusing or thrilling, just as he simply finds certain tastes appealing; his imagination is so constituted that these things appeal to him, and having found them so he forms the desire to witness them again. The pleasure in question may not be a bodily pleasure, and its source may not be the constitution of his body and its organs; it is nonetheless a brute fact about his way of being affected by the physical world that looking at corpses gives him pleasure, so that if imagination, and not the bodily organs, is its source, still, the desire for that pleasure is independent of reason's desires to know the truth and to rule his life. Whatever precisely the imagination may be it is on Plato's view linked essentially to the world as it *appears* rather than to reason, understood, as he understands it, as devoted to knowing, and governing in accordance with, the truth. This suggests the possibility of taking Leontius' castigation of his eyes as implying not that these bodily organs, but rather, more generally, that attending to the physical world independently of the discipline of reason, is the source of his pleasure. Certainly, reference to vision and its organs, the eyes, often does play this symbolic role in the *Republic*.<sup>14</sup>

What then about the democratic man's pleasure in dabbling at philosophy? If this is to be construed as an appetitive pleasure, then it must be sharply distinguished from the corresponding pleasure of the true philosopher, since that is a pleasure of reason. In enjoying philosophizing the philosopher is enjoying the pursuit of the truth; his desire for this pleasure is the expression of his reason's desire to know the truth. The democrat, then, is not led by an interest in the truth to engage in philosophical activity. What does lead him to it? Presumably, he simply finds something appealing about it: the manipulation of words, the process of deduction, the surprise of discovery, or whatever, interests and amuses him. Yet since this is unconnected with any serious pursuit of the truth, philosophy remains only a game—so it is no surprise that, on Socrates' account, the democrat only intermittently plays at it and does not acquire any deeper and more permanent attachment. His desire to philosophize, then, counts as an appetite because he attends only to the superficial, "visible" aspects of philosophy, features of it that he happens to find in-

teresting. This interest is for him as much a brute fact about his interaction with the physical world as Leontius' interest in viewing corpses. Neither of these is a recurrent biological urge, nor even such an urge modified by the addition of likings for particular tastes or smells or bodily feelings. Yet they have their ultimate origin simply in facts of experience, in the fact that the person in question happens to get a certain pleasure from doing these things, and this justifies classifying them together with the bodily appetites. They are independent of reason in the same sort of way, and can be opposed by reason on the same sort of grounds.

#### IV

We come at last to spirit: Socrates' standard name for the source of this third type of motivation, *τὸ Θυμοειδές*, derives from a Greek word, *θυμός*, that by Plato's time seems to have been in ordinary use mostly as a name for anger: the word is in fact etymologically the same as our word "fume"—someone in a state of *θυμός* would be "fuming" about something. But in Homer, where it appears very frequently, the word has a broader usage: it names the part of themselves to which Homeric heroes speak, or which speaks to them, when they are aroused for action, and into which they, or some tutelary deity, pour might and strength when their prowess is about to be put to the test. It is thus the immediate source of action, especial vigorous action, and the seat of emotion, especially those emotions (anger, for example, but also on occasion sexual passion as well) that motivate vigorous and bold action.<sup>15</sup> As we shall see, Plato's theory of *θυμός* is obviously much indebted to Homer; taking his account of *θυμός* altogether, the developments in Books VIII and IX together with the initial argument in Book IV, *θυμός* seems closely connected in Plato's eyes, as in Homer's, with vigorous, competitive action. But in his account in Book IV Plato appeals exclusively to various forms of anger, and not to any of the other desires and emotions that get assigned to *θυμός* in Homer. His examples cover a fairly wide range: they range from the fury of screaming infants (441a7-9) and barking dogs (b2-3), to Odysseus' outrage at the sexual misbehavior of Penelope's maids with her suitors (441b4-c2), to Leontius' annoyance and disgust with himself for giving in to his ghoulish fascination for corpses and the aroused sense of justice which causes a man to insist on his rights, even though the effort may cost him such deprivation and pain

as to seem hardly worth it. Later in the fourth book courage emerges as the specific virtue of this part of the soul (442b5-c3), and in Books VIII-IX it is constantly described as the honor-loving (*φιλότιμον*) and victory-loving (*φιλόνικον*) part, because, as Socrates says in one place (581a9-10), spirit is "always wholly striving for power and victory and good repute"—i.e., apparently, the reputation for effectiveness, single-mindedness, strength of character, and other "executive" virtues. (A reputation for sensitivity and compassion, or wittiness, or brains, or even judiciousness would not count in *this* context as good repute.) So the spirited part expresses itself first of all in ordinary anger of various sorts; secondly in the moral feelings of shame, outrage, and the offended sense of justice; and thirdly, in the desire to assert oneself, to be effective both in one's own private life and in the community's. What is it that in Plato's eyes links these things together—what is the principle of unity here?—and why does he think that together they constitute a third sort of motivation, coordinate with desires of reason and appetitive desires?

As before, his argument depends upon appeal to the fact of conflict. But his method of arguing from striking examples fails him in this instance. He argues first that *θυμός* is distinct from appetite, by the example of Leontius, who becomes angry at himself (more specifically at his appetite for corpse-gazing)—here appetite is opposed by anger, so this anger is a desire deriving from another source than appetite. Then he argues that *θυμός* is distinct from reason: first of all because babies and animals get furious but do not have the power to reason (that is, the power to figure out the truth of things and direct their lives in accordance with the truth), and secondly by the example of Odysseus, whose anger (more specifically, outrage) at the maids is opposed by his reason. Odysseus' anger impels him to punish the maids on the spot, but that would upset his rational plan to kill off the suitors, so his rational desire to do the latter opposes both the action proposed by anger and the anger itself.<sup>16</sup> His anger is therefore a desire deriving from another source than reason. One trouble with this two-stage argument is that it presupposes that all the cases of anger in question are of the same type and derive from the same internal source; but it is not obvious, and certainly requires argument to show, that that is so. One cannot assume just because all these phenomena can be called "anger" that they are in relevant respects all alike. Indeed, it is clear that whatever screaming babies and attacking dogs are feeling is very different from what Leontius feels, and the fact that the latter is no appetite does not imply that the former cannot be. It is conceivable, I think, that all we have here is the opposition be-

tween reason and appetite all over again—Leontius' anger being a second desire of reason opposing that ghoulish appetite, the baby's and the animal's fury and Odysseus' outrage being nothing but appetitive desires opposed, in the latter case, by reason.<sup>17</sup> What's required, in order to close this gap in Plato's argument, is a closer consideration of how these forms of anger are actually constituted, to see whether they, or any of them, really are a new kind of motivation coming into conflict sometimes with desires of reason and sometimes with appetites.

It will help, I think, in doing this, if we turn first to consider what Plato says about spirit in Books VIII and IX; we can apply what we learn there so as to eke out a satisfactory interpretation of the Book IV examples. In Books VIII and IX Socrates develops an account of four kinds of persons who lack the virtue of justice as he has defined it, because reason is not in control of their lives. We have seen already that for reason to be in control of a person's life is for him to have worked out on exclusively rational and theoretical grounds what goals are worth pursuing and to have patterned his life around the pursuit of those goals. The four bad kinds of person Socrates describes are conceived by him as people in whom another part of the soul has grown strong, displacing reason and establishing its own control over them and their lives.<sup>18</sup> So the person whom he calls timocratic is someone in whom  $\vartheta\mu\mu\acute{o}\varsigma$  has fixed the goals around which he has patterned life. Socrates describes the timocratic person as “somewhat self-willed and a little bit on the uncultured side . . . harsh with his slaves . . . but gentle with free men and very obedient to authorities, a seeker after public office and public esteem, not thinking himself worthy of office because of his ability as a speaker or anything like that, but because of his accomplishments in battle and military affairs, and a devotee of athletics and hunting” (548e4-549a7); he will be disdainful of money and the pursuit of it, while nonetheless placing a rather high value on having it (549a9-b2, cf. 548a5-b2). In sum, the person dominated by  $\vartheta\mu\mu\acute{o}\varsigma$  is “a haughty man and a seeker of public esteem” ( $\bar{\nu}\varphi\bar{\eta}\lambda\bar{\delta}\varphi\bar{\omega}\bar{\nu}\ \tau\bar{\epsilon}\ \chi\bar{\alpha}\bar{\iota}\ \varphi\bar{\iota}\bar{\lambda}\bar{\o}\bar{\tau}\bar{\iota}\bar{\mu}\bar{\o}\bar{\varsigma}\ \bar{\alpha}\bar{\eta}\bar{\bar{\eta}}\bar{\rho}$ , 550b7). It is worth emphasizing that Socrates only claims that this kind of outlook results where  $\vartheta\mu\mu\acute{o}\varsigma$ -motivations are not only particularly strong in a person but develop freely, without being trained and directed in subordination to other values: he insists that the people who will make the best warriors in his ideal republic must be by nature unusually “high-spirited” ( $\vartheta\mu\mu\acute{o}\varsigma\bar{\epsilon}\bar{\iota}\bar{\varsigma}\ 375a11-12$ , e10), but the description just quoted will not fully apply to them because they have been educated to respect philosophical values and to seek the good of their fellow citizens, so that though  $\vartheta\mu\mu\acute{o}\varsigma$  gov-

erns them *what*  $\vartheta\mu\mu\acute{o}\varsigma$  directs them to do will not be the same as for Socrates' timocratic man. And, of course, where  $\vartheta\mu\mu\acute{o}\varsigma$  is subordinated to appetite, as in the person whom Socrates calls oligarchic, it will bring its special motivations in support of the agent's dominant appetitive values: the oligarchic man does not permit his  $\vartheta\mu\mu\acute{o}\varsigma$  to “admire and esteem ( $\tau\bar{\iota}\bar{\mu}\bar{\alpha}\bar{\nu}$ ) anything else but wealth and the wealthy or to seek public esteem on any other ground than the possession of money and whatever else contributes to that” (553d 4-7).

The central idea suggested by these and other passages of Book VIII is that  $\vartheta\mu\mu\acute{o}\varsigma$  is understood by Plato as that wherein one feels (a) the competitive drive to distinguish oneself from the run-of-the-mill person, to do and be something noteworthy within the context provided by one's society and its scheme of values; (b) pride in oneself and one's accomplishments, to the extent that one succeeds in this effort; (c) esteem for noteworthy others and (especially) the desire to be esteemed by others and by oneself. Because competitiveness can be so variously directed, and the bases of self-esteem (and pride and esteem for others) can vary so widely,  $\vartheta\mu\mu\acute{o}\varsigma$ , if this is what it is, can in different people support widely different courses of action and ways of life, and this Plato claims it does. But it does not seem to me unnatural to think that someone in whom competitiveness and the desire for esteem and self-esteem were particularly strong should tend toward the athletic, military, and political pursuits, to which Plato says the  $\vartheta\mu\mu\acute{o}\varsigma$ -dominated person will especially devote himself; these are obvious, as well as traditional, activities in which a man, at any rate, can hope to make himself stand out from others as esteem and self-esteem require and competitiveness implies.

I suggest, then, that the motivations that Plato classifies under the heading of spirit are to be understood as having their root in competitiveness and the desire for self-esteem and (as a normal presupposition of this) esteem by others. Can we make sense of Plato's examples of anger in Book IV along these lines? Three of the five fall immediately into place. When Odysseus in disguise comes upon Penelope's maids cavorting with her suitors his immediate impulse is to punish them on the spot: the sight of such disorder in his own household is naturally a blow to his self-esteem (self-respecting noblemen don't permit that kind of thing), and his anger is a response to this affront. It urges him to act immediately to restore order and therewith prove himself deserving of the esteem which he feels is placed in jeopardy by the continuance of this state of affairs. His anger thus represents a traditional view of things to which his continued self-esteem is tied: he will feel bad about himself

unless he acts at once to vindicate his honor. Yet his reason does not support this traditional view: from reason's point of view delay does not mean indifference or weakness or cowardly acquiescence, and there is (Odysseus thinks) no *reason* for him to think less well of himself for delaying (in fact, quite the contrary, since he plans eventually *both* to punish the maids *and* to kill off the suitors). But though this is how he *thinks*, it is not how he *feels*. The reaction of his θυμός shows that his self-esteem, the way he feels about himself, is tied up with a certain traditional view of the king's dignity, not with the view implied by his own rational planning. Hence reason and spirit in his case are in conflict over what to do. A bad upbringing, Socrates suggests (cf. 441a3), has corrupted Odysseus' spirit, causing him to feel differently about things than he thinks.

Similarly for Leontius. On a considered view of things Leontius rejects corpse-gazing as a bad thing or at any rate nothing to take any interest in. Yet he continues to have an appetite for that sort of thing. Unlike Odysseus, Leontius' θυμός is in agreement with his reason: he feels that corpse-gazing is *sordid*, and does not want to be the kind of person who goes in for it; in fact, perhaps, he aspires to be the kind of person who makes the goals of reason his goals and has no others. Hence when he incontinently acts on this rejected desire not only does his reason disapprove of what he has done, but he also suffers a blow to his self-esteem: the anger he feels at himself (it might equally have been shame or simply exasperation) is the natural response to this failure to measure up in his own eyes. The situation is the same with the man who responds with anger to what he judges is unjust treatment: it is natural to think that the perceived injustice is taken by him as a sign that the perpetrator disregards or belittles him and his interests, and his anger is the normal and natural response to such a slight. Not to become angry would be a sign that one acquiesced in the perpetrator's estimation of one's worth or importance, and no one who feels self-esteem could do that. So here too anger expresses the competitive desire to acquire and preserve self-esteem.

The other two examples of θυμός appealed to in the Book IV argument are less easy to accommodate. Screaming two-week-old babies and ferocious dogs presumably have no self-conception (I assume the dogs are not even self-conscious) and so though their anger may express some primitive form of competitiveness, it is at any rate not a form that has anything to do with self-esteem that their anger expresses. But perhaps Plato counts these cases of anger as motivations of the same kind as Le-

ontius' and Odysseus' because he sees them as the central primitive phenomena which get transformed, as we mature, into the full-fledged competitive desire for self-esteem that expresses itself partly in anger like Leontius' and Odysseus', as well as in the admiration and emulation of others, the disdain for anything lowly, and the aspiration for solid accomplishments which we have found attributed to θυμός in Books VIII and IX.

If I am right that competitiveness and the desire for esteem and self-esteem lie at the center of what Plato understands by θυμός in the *Republic*, it is not difficult to show that θυμός-motivation is a different kind of motivation from either the desires of appetite or the desires of reason, as Plato construes them. It is different from appetite because appetites lack the self-reference which is essential to esteem and self-esteem; and it is different from the desires of reason, which may of course be self-referential, because of the way in which it is constituted. There are two aspects to this difference between θυμός and reason. What a θυμός-desire desires is competitive success and the esteem from others and oneself that comes with it. Like all objects of desire one can, of course, say that in desiring all this, θυμός (or the person *qua* experiencing θυμός) regards it as good. But that does not mean that a θυμός-desire is a desire for good (a "good-dependent" desire) in the way that the desires of reason are. The difference has to do, so to speak, with the order of priority between the desire itself and thoughts about good. In the case of reason, thoughts about what is good come first, a desire being formed for whatever one thinks (rightly or wrongly) is good. But in the case of θυμός the desire for competition and esteem comes first (without regard to any antecedent question whether these things really are good, or if they are why they are so), and thoughts about good then follow.

This difference connects with a second one. For although as Plato says a person's θυμός tends by nature to support his reason's judgments about good and desires for that, it does not always do so. That is because the *origins* of one's θυμός-desires do not in any event lie in rational processes of reflection, but in all kinds of contingencies in one's upbringing and subsequent life. Θυμός develops under the influence of how other people (especially one's parents) respond to and treat one. How we feel about ourselves—under what circumstances we experience a blow to our self-esteem, what we aspire to be and do, what competitions we enter—are to a large degree determined by our experiences in childhood, even if as adults we can partly remove or refine the effects of

our upbringing so as to make the way we feel about ourselves conform with our rational conceptions of how we ought to live. It is possible, even normal, to find oneself, as Odysseus did, with conflicting conceptions and attitudes, some derived from the influence of events in earlier life in forming the basis of our self-esteem, others the product of considered rational judgment. It is natural, Socrates says, for these attitudes to be in agreement, for a person to feel good and bad about himself in just the ways that conform to his rational view of how he ought to live.<sup>19</sup> This is the result of the inherent authority of the truth, which is ideally the possession of reason, on which both his rational view of things and the basis of his self-esteem ought to converge. But they do not always do so, and even when they do the basis of a person's self-esteem is to be accounted for not simply by appeal to any rational argumentation he went through but to his personal experience in his developing social relationships.

On Plato's tripartite theory, then, competitiveness and the desire for esteem and self-esteem are an innate form of human motivation, distinct from the appetites and reason itself and equally as basic as they are to human nature. There is certainly no denying that this kind of motivation, in its many guises, does play a very large part in the conduct of any human being. Any plausible theory of human motivation must surely pay special attention to it. It is a considerable merit of Plato's theory of the human soul in the *Republic*, whatever its other shortcomings, that it gives fuller and more explicit recognition to this fact than subsequent theories have done.<sup>20</sup>

## Notes

1. In referring in this paragraph to the conception of virtue espoused by the character Socrates in Plato's early dialogues as "Socratic" I follow the by now conventional scholarly practice, according to which this character's central views are attributed to the historical Socrates. Conventions are dangerous things, and this one should certainly not be accepted as uncritically as it often seems to be (for a recent defense of it, see W. K. C. Guthrie, *A History of Greek Philosophy* III [Cambridge: Cambridge University Press, 1969], 349-55). It is worth emphasizing in this connection that, though of course he also had other evidence now lost to us to go on (the oral tradition, plus writings of Antisthenes, Aeschines, and other Socratics), Aristotle plainly attributes to the historical Socrates essentially the same views on the virtues that one finds Plato's character Socrates espousing in the early dialogues (see esp. *MM* I 1,1182a15-23; also, *MM* I

20, 1190b28-32; 34, 1198a10-12; *EN* VI 13, 1144b17-21, 28-30; III 8, 1116b3-5; *EE* I 5, 1216b3-8); and at *EN* VII 2,1145b22-27 and 3, 1147b14-17 verbal echoes with the *Protagoras* (compare 352b8-c2) strongly suggest that he relied directly on Plato's dialogues at least some of the time for his conception of the historical Socrates' philosophical views. So Aristotle's treatment of Socrates confirms the correctness of this convention, however antecedently dubious it might seem.

2. Thus Aristotle (*EN* I 13, II 5-6) describes virtue of character simply as the proper coordination between reason on the one side and nonrational desire, in general, on the other. He says nothing in this context about any differences there may be in kinds of nonrational desire. In other parts of his ethical theory, however, Aristotle does in fact preserve the distinctions that led Plato to regard the human soul as having three parts. He regularly divides ὄρεξις (desire) into three subkinds, βούλησις, θυμός, and ἐπιθυμία (see *de An* II 3, 414b2, III 9, 432b3-7; *de Motu* 6, 700b22; *EE* II 7, 1223a26-7, 10, 1225b25-6; *MM* I 12, 1187b36-7), and he assigns the first to reason itself (*de An.* 432b5, 433a23-25; *Top.* IV 5, 126a13), making the latter two belong to the "nonrational element" (*de An.* 432b6). Thus Aristotle holds (with Plato; see below) that reason has a special kind of desires of its own and he divides nonrational desires into the same two species as Plato recognized (see below). His acceptance of the Platonic theory that there are three distinct kinds of desire has important though frequently unappreciated effects on his moral psychology, as can be seen for instance in his concept of προσλήψεις (decision, rational choice): *EE* II 10 makes clear (see 1226b2-5, 1227a3-5), as *EN* III 2-3 does not, that the ὄρεξις that is according to Aristotle a component of προσλήψεις is a βούλησις i.e., a desire belonging to reason itself, and not any nonrational desire. (J. Burnet, presumably relying on these *EE* passages, attributes this view, correctly in my opinion, to Aristotle in the *EN* too: βούλησις, he says in commenting on *EN* III 3, is "the appetitive element in προσλήψεις," *The Ethics of Aristotle* [London: Methuen, 1900], 109, 131, 132.)

3. See most recently T. Penner, "Thought and Desire in Plato," in G. Vlastos, *Plato II* (New York: Anchor Books, 1971), 111-13; also W. F. R. Hardie, *A Study in Plato* (Oxford: Clarendon, 1936), 142-43, and F. M. Cornford, "Psychology and Social Structure in the *Republic*," *Classical Quarterly* 6 (1912): 262-64.

4. This is certainly suggested by his remark at 435e-436a that if, as the foregoing political analysis has asserted, there are three types of persons suited for three distinguishable kinds of social work, that can only be because there are in each human being three psychological elements or powers, the special strength of one or another of which in a person is what makes him belong to one or another of the three social types. Similarly, at 544d6-e5 (cf. 545d1-3), Socrates argues that what determines the character of a city as timocratic, or oligarchic, or democratic, etc., is the character of those individuals in it who are in

command: where people dominated by spirit, concerned about competitive values, govern, the city will be a timocracy (547el-4, 548c6-7), and so on for the other cases.

5. See for example *Leviathan* ch. 8: "For the thoughts are to the desires as scouts and spies, to range abroad and find the way to the things desired." Bernard Gert argues (Introduction to Hobbes, *Man and Citizen* [Garden City, N.Y.: 1972], 13-16), that Hobbes does not limit reason to this scouting and spying function, but thinks that in addition it seeks *one* end not set by passion, viz. the avoidance of one's own violent destruction. On Gert's view not Hobbes but Hume is the originator of this modern view. It seems best, however, to interpret Hobbes as holding that the avoidance of violent destruction is the object of a settled and constant passion experienced by all persons that serves as a background against which varying particular passions arise and decline. On this interpretation Hume's conception of reason as only "the slave of the passions" (*Treatise*, ed. Selby-Bigge, 415) is just a reformulation of the Hobbesian view; Hobbes deserves the credit or blame for originating the familiar modern view.

6. I have been anticipated in this interpretation by R. C. Cross and A. D. Woozley (*Plato's Republic: A Philosophical Commentary* [London and New York: 1964], 118-19).

7. And cf. 554d-e, where Socrates speaks of a conflict among appetites in the "oligarchic" man's soul; his "better" appetites (his love of money, thriftiness etc.) do battle with and win out over his "worse" ones (his occasional extravagant lusts, thirsts, hungers). This man, Socrates implies, has a kind of self-control, but one that is far from being a virtue, since the appetite that prevails keeps control not (as reason would do) by the logical and rational force of ideas, but by inducing instinctual and irrational fear—the irrational fear of what will happen if money is spent in order to gratify the base appetites. Plato shows no sign of discomfort here in recognizing conflicts within what he continues to think of as a single part of the soul. This is reasonable if he did not mean to argue in Book IV that just any conflict of desires betrays a difference of origin (i.e., a difference of type of motivation) in the desires, but hardly otherwise.

8. *Plato's Moral Theory* (Oxford: Clarendon, 1977), 327.

9. Here I mean to attribute to Plato the stronger of two possible models one might have in mind for what it is for reason to rule in our lives. (1) According to the weaker view reason as a ruler accepts desire as the ultimate criterion of value; on this view, that a thing is, or would under certain conditions come to be, desired (whatever the nature and source of the desire in question) is reason's sole basis for assigning actual or potential value to anything and so giving it a weight in its calculations. Given this criterion of value, and the facts about what one desires or might come to desire, reason's role is to work out a best overall scheme of life, with strategies and tactics for dealing with particular problems that may arise, and to decide on appropriate action in individual circumstances. In carrying out this task, reason aims at satisfying one's desires as fully as possible, taking into account how much one wants various things, how distressed

one would be without them, how getting or failing to get something one wants affects one's ability to get or enjoy other things one wants, and so on. On this model, for reason to rule is (a) for it to be free to decide, upon an impartial survey of the relevant facts about the world and about one's desires, how one should live and act, and (b) for its decisions to be effective. Once reason decides on an object of pursuit or a mode of action in some situation it may contribute a new desire of its own (the desire to pursue or do that *because* it is supported by reason), but this desire only comes in as a reinforcement of the antecedent desires whose satisfaction reason was previously deliberating about. (2) On the second, stronger model, reason's work, and the desires it gives rise to, are more fundamental. Here, instead of taking desire as the criterion of value in its object, reason presumes to be able to decide by appeal to its own principles what things are good and how good they are; that, as may happen, these are also desired, and the degree to which they are desired, have nothing to do with their value (except to the extent that having a desire may constitute recognition of some antecedent value). It would not be easy to specify what according to Plato these principles might be, but the following example may indicate the general idea. We speak of the good of living things in general (not just animals), and we consider a creature's good to consist, at least in part, in its attaining, and functioning in, its natural mature state. The satisfaction of desire obviously cannot be the basis for such a judgment where plants are concerned, and it is not implausible to exclude it even in the case of animals. We might well expect an animal to find satisfaction for its desires in the natural functioning of its mature state, but if it did not one need not conclude that its good lay elsewhere, but only that, through some perversion, it failed to enjoy its good when it had it. In the *Republic*'s theory the function of the form of the Good is to provide the knowledge of those principles of goodness that will permit reason to work out a scheme of ends for an individual to aim at achieving in his life and to make particular decisions as circumstances require (on this see my "The Psychology of Justice in Plato," *American Philosophical Quarterly* Vol. 14 [1977], 151-57). Given this knowledge one will know, for example, that and why eating or drinking is a good when it is (because health requires it, and health is a good); but from reason's point of view one's appetites for food and drink themselves provide no reason at all for thinking that these are good things. Similarly for all other nonrational desires.

10. This interpretation of Plato makes it easy to see how Aristotle might have arrived at his distinction between theoretical and practical reason. Theoretical reason is simply reason used to pursue one of the two ends that according to Plato rational beings qua rational have got, viz. to know the truth; practical reason is reason pursuing its other end, the end of ruling our lives. Hence Aristotle can say that in both employments reason aims at truth (*EN* VI 2, 1139b12)—not truth in the one case and something else (say, good) in the other. For he, just as much as Plato, conceives of reason as having the power to rule in the stronger of the two senses distinguished above (n. 9), and accordingly the desire of reason to rule is for him the desire to achieve and enforce practical

truth, i.e., the correctness of ends as well as means.

11. And notice that at 437d11-e2 “modified thirst,” while Socrates insists it is not merely *thirst*, is nonetheless classed as an ἐπιθυμία.

12. Plato justifies these epithets at 580e5-581a1 by saying that the principal use of money is to provide the means by which the appetites can be gratified.

13. In developing his account of the various types of unjust person (timocratic, oligarchic, democratic, tyrannical) Plato makes it clear that, just as the “timocrat” (550b5-6) has yielded the governance of his soul to his Θυμός, so the “oligarch,” the “democrat,” and the “tyrant” are all ruled in different ways by appetite. The “oligarch” is said explicitly (553c4-7) to enthrone appetite as his ruler, which means that in his plans and decisions his ultimate aim is constantly and only to gratify appetite; being dominated by appetites, he forces the other parts of the soul to want and get satisfaction only from assisting in this effort (553d1-7). But his pursuit of appetite takes the perverted form of aiming at fulfilling first what Plato calls the “necessary” appetites and, beyond them, only the appetite for the mere accumulation of the *means* (money) whereby these and other appetites might be satisfied. The “democrat” (559d-561e) refuses to follow the oligarch in repressing his other appetites, and ends by establishing equality for all appetites: he allows himself to acquire any and every appetite that his circumstances and nature make it possible for him to acquire, and then he indulges all his appetites in turn, on an equal basis. A consequence of this account is that, as noted in the text, when, in accordance with his scheme, the democrat goes in for athletic, political, and philosophical pursuits, the desires for these things that he is bent on indulging must be construed as appetites, desires belonging to the ἐπιθυμητικόν, odd as this may seem. They are not desires of spirit or reason.

14. The ὄρατὸς τόπος (508c2) described in the analogy of the sun is also the realm of τὸ γιγνόμενον τε καὶ ἀπολλύμενον, which if the soul attends to in working out its general conception of things, instead of to the νοητόν, it will fail to reach any understanding (d6-9); and when a soul does that it is reduced to taking resemblances for reality (476c5-7) and ruling in accord with the false and inadequate conventional standards (479a3, d3-5) that have been developed over time by others who likewise relied on experience without philosophical thought to guide their lives.

15. David B. Claus in the latest discussion of the Homeric usage of soul-words, *Toward the Soul* (New Haven: Yale University Press, 1981) argues that in Homer Θυμός like μένος, ἥτορ, and κῆρ, with each of which it is in many contexts easily interchangeable, has the central meaning of “life-force,” but with a special connection to personal affection (see his ch. 1, esp. 37-42).

16. Notice that in the passage Plato cites from Homer (*Od. XX, 17; Rep. 441b6*) Odysseus addresses and reproves his κραδίη, i.e., his anger or his heart conceived as the seat of it, bidding it to be calm and endure without making a fuss. The conflict in this case, as in that of Leontius where spirit and appetite

conflict, involves a direct criticism by the “higher” faculty of the “lower,” and not merely conflicting impulses to action.

17. This is the view taken by Cornford, Hardie, and Penner (see locc. citt. n. 3 above).

18. Plato's highly metaphorical talk of the displacement of reason from its throne and the usurpation of power by Θυμός or by appetite (550b4-6, 553c4-d7) is potentially very misleading. He does not mean either that reason ceases altogether to function (see 553d) or that the usurper actually begins to perform reason's functions of calculating out what to do, declaring where the overall good lies and deciding accordingly. What happens instead is simply that, yielding to the importunities of the usurping desires (i.e., accepting the strength or frequency of these desires as criterion of the value of their objects), the person's reason comes to adopt, as its *own* general view of what is good for him, the overall plan of gratifying those desires first and foremost. In doing this reason fails to perform its most essential task, namely to work out on its *own* theoretical grounds where the good actually lies; and that is why Plato says reason is no longer ruling in such a person's life, and why he says that instead those other desires rule, the ones to which reason has abdicated its own responsibility to set goals. But reason continues to be the only part of the soul in which judgments about overall good and those desires for good that follow upon them are located.

19. Thus Socrates says (441a2-3), understandably enough, that Θυμός is by nature reason's helper (ἐπίχουρος), and he describes it as entering disputes between reason and appetite or reason and outside agents as the ally (σύμμαχος) of reason (440b2-4, c7-9). It is harder to understand why Socrates so flatly insists (440b4-7) that one *never* finds anyone's Θυμός intervening in a dispute between reason and appetite on appetite's side. He himself later describes the oligarchic man as enslaving both his Θυμός and his reason to his appetite for money (553d), but just as the dominance of this appetite is not sufficient to prevent spendthrift desires from arising altogether (554b7-c2), so it presumably is not sufficient to prevent reason or Θυμός from occasionally rebelling and generating desires not subordinated to the pursuit of wealth. And if, after all, reason and Θυμός are independent of one another, why could not a desire of reason (e.g., to spend some money for the public good) arise that conflicts with the master-appetite, only to be opposed by Θυμός for that very reason? When, as with the oligarchic person, Θυμός has been habituated to support appetite, this is only what one should expect: reason's desire to act generously should be felt by Θυμός as disgraceful—soft-hearted, a sign of weakness or sentimentality, etc. That Socrates does not envisage this possibility in Bk. IV is presumably to be explained by supposing that he assumes there that appetitive gratification is such a simple and easy thing to arrange, or if difficult so obviously in itself nothing to be proud of, that when reason opposes it the love of competition *could* not find any scope for activity except on reason's side (i.e., on the side whose winning out might show that something worth crowing over had been achieved). This

thought seems natural enough, and appropriate to the context in Book IV; but it is nonetheless quite naive, as the account of the oligarchic man in Book VIII shows.

20. I am grateful for helpful comments on earlier versions of this paper from a number of people, especially Annette Baier, Gail Fine (my commentator when I presented the first version at Cornell), Cynthia Freeland, John Hare (commentator on a version delivered at a New Jersey Regional Philosophy Conference), and Alexander Nehamas.

## 5

### Plato's Analogy of Soul and State

Nicholas D. Smith

At the beginning of Book II of Plato's *Republic*, Glaucon and Adeimantus express dissatisfaction with Socrates' refutation of Thrasymachus, and challenge him to offer a more persuasive proof that justice is preferable to injustice. But at the end of Book I, Socrates compared the conversation with which Glaucon and Adeimantus are so unimpressed to a feast of gluttons, where the rude diners attempt to devour each new dish before properly enjoying the preceding one (354al2-b5). To have conducted the inquiry properly, Socrates proclaimed at the end of Book I, would have been to discern first what justice is before turning to the next "dish" and attempting to discern whether it is preferable to injustice or not (354b5-c1).

In giving his reply to the renewed Thrasymachean challenge mounted by Plato's brothers, then, Socrates ensures that there will be a proper "feast" by turning first to the question of what justice is, before any attempt is made to decide whether justice or injustice is preferable. Famously (or notoriously), to accomplish this goal, Socrates proposes the analogy of large and small letters (368c7-d7), in which the nearsighted are given the opportunity to read the same message in large letters, before attempting to discern it written in small letters. This situation, in turn, is compared to the attempt Socrates proposes, to "see" justice first in a state, where it will be "writ large," before trying to figure out what it is "in smaller letters" in a soul.

Plato's Socrates and his interlocutors accept from the outset and without argument that justice in the state and soul are like the same letters inscribed in a large and a small format, and this assumption has been widely challenged and questioned by Plato's critics.<sup>1</sup> In this paper, I will neither object to this assumption nor defend it, but instead will show that the specific way in which Plato articulates his analogy, and the way in which this analogy is generally supposed by scholars to work—if it works at all—is irreparably logically flawed.<sup>2</sup> Specifically, I will argue in Part I that the arguments Plato offers for the tripartition of the soul are founded upon an equivocation, and that each of the valid options by which Plato might remove the equivocation will not produce a tripartite soul. The result of this argument, plainly, is that the way in which Plato actually makes the analogy of state and soul—by appealing to the similarities between a tripartite state of rulers, warriors, and craftsmen, in the *Kallipolis*, to a tripartite soul of reason, high spirit (or *thumos*), and appetite—cannot validly be made. (In the remainder of this paper, I will call this specification of the soul-state analogy the “3-3 specification”). In the second part of my argument, however, I will argue that Plato reveals a lack of full commitment to the 3-3 specification of the analogy without thereby calling the analogy itself into question. If this is so, then it follows that the heart of the analogy is not to be found in the comparison of the *Kallipolis* and its three parts to the soul conceived as tripartite, but rather must be supposed to reside in some other connection between the ways in which justice characterizes states and souls, and I will suggest what this other connection consists in. I will conclude by showing that the failure of Plato's psychological arguments that I find in Part I of this paper, which is surely fatal to the 3-3 specification of the analogy, is not fatal to the analogy itself, or even to the political aspects of Plato's *Republic*, even if it would require that substantial changes be made in Plato's psychological theory.

### I. Plato's Arguments for Psychic Division

At 434b9-d1, Socrates and Glaucon agree that they have discovered what justice is in the city:

For the three of these classes to interfere in one another's business or to exchange with each other, then, is the worst damage to a city and might rightly be called the worst evil.  
Indeed.

And don't you call injustice the worst evil to one's city?  
Of course.

So this, then, is injustice. Moreover, to say it the other way around, for the wealth-making, auxiliary, and guardian classes to mind their own business and do their own in the city, is the opposite—justice—and is what makes the city just.

It is this and nothing else, I agree, he said.

The conception of justice expressed in this passage links it directly to the tripartite class structure of the *Kallipolis*, and so it is not surprising that when Plato's Socrates goes on to consider what justice is in the soul, we find him immediately looking for three parts of the soul to correspond to these three classes in the *Kallipolis*.

The city was believed to be just when the three natural classes within it each did its own, and it was temperate and courageous and wise because of other conditions and states of these classes.

True, he said.

And so, my friend, we will evaluate an individual as rightly called by the same names as we give in evaluating a city, if he has these same forms in his soul, and the same conditions in them.

Quite necessarily, he said. (435b4-c3)

The general principle, which Socrates employs in each of the partitioning arguments, is announced at 436b8-c1:<sup>3</sup>

Plainly, the same thing cannot do or undergo opposites in the same respect<sup>4</sup> and in relation to the same thing, at the same time. So, if this should ever happen in these things, we will realize that they weren't the same, but many.

Socrates immediately goes on to introduce a paradigm case for the application of this principle: the same thing at the same time in the same part of itself cannot both be at rest and also move (436c5-6). From the way in which he articulates the principle, and immediately clarifies his meaning with this example, we might well understand Socrates as committing himself to nothing more controversial than a principle of non-contradiction:

(X) The same simple object cannot at the same time and with regard to the same things perform contradictory actions or undergo contradictory changes.

This interpretation of the principle is especially inviting if we construe moving and resting as contradictories—that is, where either “resting” *just is* “not moving” or “moving” *just is* “not-resting.” But Socrates actually states his principle in a way that invites an alternative understanding, by stipulating that it is not *opposite* actions or modifications that a thing cannot do or undergo, where by “opposite,” we do not mean “contradictory.” Pushing and pulling, for example, are not contradictories, but opposites, and this may be the sort of model that Plato has in mind.<sup>5</sup>

**(O)** The same simple object cannot at the same time and with regard to the same things perform opposing actions or undergo opposing changes.

Evidence that **(O)** and not **(X)** should be our interpretation of the partitioning principle may be seen from Socrates’ subsequent list of examples of the sorts of “opposites” that he has in mind: “assent and dissent, desiring to have something and rejecting it, drawing in and thrusting away” (437bl-5). This list, however, requires interpretation, as we shall soon see.

The problem comes when Plato seeks to apply his partitioning principle to the specific cases he gives to divide the soul, for the relevant cases either do not appear to fall under the above interpretations of the principle—or if they do, the principle will turn out to do considerably more than Plato seems to want it to do. To see that this is so, let us begin by reviewing Plato’s first partitioning argument, which is intended to show that the soul may be divided into at least two parts: reason and appetite.

At 439a1-e1, Plato considers the case of someone who is thirsty, but refuses to drink. But this would be a case to which we could apply **(X)** only if it would be true to say of the person both that she did and also that she did not wish to drink. As tempting as it might be to say this about her, it would not be true, however. We often do describe our *negative* or *aversive* desires in this way: “I do not want to drink,” but it is important to distinguish between cases in which we wish to proclaim a *lack* of a certain desire (which would be the contradictory of *having* the relevant desire) and cases in which we have a *negative* desire, or an *aversion* to something, for example, where we have very real desire *not* to drink. Non-sentient entities *lack* any desire of any sort; so it is true to say of a thumbtack, for example, that it does not want to drink. It is most certainly *not* true, however (assuming my assessment of thumbtacks as non-

sentients is correct, of course), ever to say that a thumbtack wants *not* to drink. Thumbtacks lack desires, but never have aversions.

What appears to be happening in Plato’s case of someone who is thirsty but refuses to drink, then, does not fall under principle **(X)** because it does not identify a case in which the person both *has* and *lacks* the relevant desire. Instead, the person seems to have *two* desires, which appear to oppose one another. Accordingly, it would appear that **(O)** is a more likely rule to apply in this case.<sup>6</sup>

One of the most notable—and perhaps also most troublesome—aspects of psychological states is their “intensionality.” The objects to which such states refer, it is often said, are given “under description.” Accordingly, it is one thing to say of Oedipus that he sometimes wished to enjoy marital relations with Jocasta, and quite another to say that he sometimes wished to enjoy marital relations with his mother. We may fairly suppose, given Sophocles’ telling of the tale, at least, that the former is true and the latter is false, about Oedipus.

Once we recognize this feature of desires, however, two distinct possible interpretations of **(O)**, as it would apply to desires, are open to us, depending upon what we are willing to count as “the same object.” In one sense—the extensional sense—the woman whom Oedipus desires is the same object as Oedipus’ mother. In another sense—the intensional sense—these are different objects. Accordingly, we can understand **(O)** to apply to cases of extensional identity, or to cases of intensional identity. Let us call these alternatives “**Oe**” for the extensional case applied to desires and “**Or**” for the intensional case applied to desires:

**(Oe)** The same part of the soul cannot be responsible for both desiring and also having an aversion to the same object at the same time, even if the object is taken, in the desire and the aversion, under different descriptions.

**(Or)** The same part of the soul cannot be responsible for both desiring and also having an aversion to the same object at the same time, taken under the same description.

Now, on the one hand, **(Or)** looks like a very plausible principle—after all, it is difficult to imagine how a person might both desire and also have an aversion to the same thing at the same time, where the thing desired and shunned was represented in both the desiring and in the shunning in the very same way. If such cases are possible at all, it seems

plausible to think that they would only be possible for someone with a wholly bifurcated desire system, as for example, we might find in someone suffering from multiple personality disorder.

It is anything but obvious, however, that this is the sort of case Plato has in mind here. He doesn't tell us all the details about why his thirsty person refuses to drink, but he does mention that the case he has in mind is one where the person has some *reason* to avoid drinking, while what makes the person desire to drink is "the result of affections of diseases" (439d1). It looks as if the case, then, is one in which the person's desire to drink takes the intended object (say) as "thirst-quenching," whereas the same object is taken by the opposed reason as "unhealthy." This, then, is more like the case with Oedipus I mentioned earlier—the attraction and the repulsion represent the relevant object in entirely different ways. If this phenomenon falls under any principle, then, it must be (*Oe*).

But as a partitioning principle, (*Oe*) is too strong, and if it is this principle that Plato wishes to employ to divide the soul, his argument threatens to atomize the soul. Plato wants it to turn out that all examples of thirst, hunger, and lust—and perhaps also all desires for material acquisition—will belong to the appetitive part of the soul. All anger, and presumably all desires for honor, will belong to the high spirit. Calculation and reasoning belong to the rational part. But cases may readily be found in which examples will fall under (*Oe*) that would require Plato to partition the soul into more than three parts.

Consider our aversion to bodily pain, for example. To which part of the soul would Plato think this belongs? Parity of reason would suggest that it should belong to the appetitive part, since the desire for bodily pleasure seems entirely to belong to this part. If this is so, then how would (*Oe*) apply to this case: Jones finds himself caught between a strong desire to drink the water he sees in front of him, on the one hand, and his aversion to the pain he anticipates if he were actually to attempt to drink the water, since he notices that it is boiling hot. He desires to drink, but holds back.

Now, perhaps Plato would count this as a case of reason opposing appetite, because it is a calculation involving anticipated harm that opposes the thirst. But surely we can imagine a case in which Plato could not count the relevant aversion as a feature of reason—for example, where Jones has sound medical evidence that the boiling water will actually serve to cure some infection in his mouth, and the burns, though painful, will cause no permanent damage. In this case, the aversion to the pain would actually work *against* Jones's reason, and surely Plato would

have to count such aversions, therefore, as opposing *both* Jones's reason and Jones's desire for drink.

Moreover, it makes no sense to suppose that such aversions must belong to the high-spirited part of the soul. Even if we tried to defend Plato's account by supposing that the high spirit was the "part" to which they belonged, we can just as easily add to our account not only that Jones has good medical reasons to overcome his aversion to the pain the boiling water would cause, but also that Jones saw that refusing to drink, in these circumstances, would count as an example of disgraceful cowardice. In such a case, we would have to understand Jones's aversion as opposed to elements of *all three* of Plato's psychic parts. Accordingly, then, if we apply (*Oe*) to this case, we must find a fourth part of the soul. Nor is there any hope that the process will stop here, for it seems obvious that in the right circumstances, literally *any* of our desires might come into conflict with *any* of our aversions, and the same kinds of scenarios—in which support for either desire or aversion can be found in reason or high spirit—will always be possible. For example, I am thirsty but to get a drink I must leave my warm bed and suffer the cold in the room;<sup>7</sup> or, I am hungry, but there is no food in the house and I can't bear the thought of fighting through traffic just to go to a store; or, I find myself sexually attracted to someone, and at the same time find myself repelled by the scent of garlic on her breath. If such examples are enough to partition the soul, to accommodate all the ways I might find myself pushed and pulled at the same time in regard to the same object-in-the-world—even if my attraction and my repulsion are not focused on the same *aspect* of that object, as it were—then it is difficult to see how the process of partitioning will not go on indefinitely. In any case, it certainly will not end after only Plato's three parts have been identified.<sup>8</sup>

## II. Justice in the State

As we have seen, the text makes it perfectly plain that Plato understood the analogy of state to soul in a way that presumes what I have called the "3-3 specification" of the analogy. If I am right about the failure of Plato's arguments for the partitioning of the soul, however, Plato does not have a proper claim to this specification of the analogy, for he has no good reasons for dividing the soul into three parts, which is required by the analogy. But the text (happily) also supplies us with compelling reason for supposing that Plato did not think that the 3-3 specification of the

analogy was required for the analogy to succeed, since he shows that he is not completely committed to either of the “threes” in this specification, in circumstances where the analogy itself is not, thereby, called into question.

Let us consider, first, the tripartite division of the state. One of the consequences of supposing that the analogy requires the 3-3 specification is plainly that no state that is not divided up into three parts could satisfy the analogy. The purpose of the analogy, recall, is to make it easier to discern what justice is, by first trying to see it “writ large” in a state, before looking for justice written in the smaller letters of a soul. If Plato supposed that this analogy required a 3-3 specification, then,<sup>9</sup> we should not expect to see him “looking for justice” in a state that is not divided into three parts. And yet, this is exactly what we find him doing—long before he has introduced the three parts of the *Kallipolis*—at 371e, when Socrates and Adeimantus have completed what Glaucon later disparages as no better than a “city of pigs” (at 372d4-5). Despite Glaucon’s disgust, however, Socrates seems prepared to “look for justice” in this very basic city.

Well then, Adeimantus, has our city grown to completeness?

Perhaps.

Where, then, in it, can justice and injustice be found? With which of what we considered did it come to be?

When Adeimantus admits that he doesn’t know the answer, Socrates exhorts him not to give up and to consider the question further (372a3-4). Even after Glaucon has shown his contempt for this city, Socrates pauses to praise this basic city one last time, before going on to consider the “fevered” city that Glaucon wishes to consider, by calling the more basic city the “true city” (372e6) and the “healthy one” (372e7).

Socrates’ willingness to look for justice within this first city—which is not only expressed but then emphasized by his exhorting the reluctant and puzzled Adeimantus and then by his remarkably positive final appraisal, to Glaucon—would be inexplicable if Plato wished to show us that justice in a city *had* to be understood in terms of the proper functioning of three classes within a city. The city that Glaucon calls a “city of pigs,” plainly, has several different functions to be performed, but all would fall within what is identified as the class of craftsmen in the *Kallipolis*. So, either there are several more than three classes in the basic city, or else there is only one class, depending upon how one wished to individuate classes. Nonetheless, Plato’s Socrates is entirely ready to

look for what it is in this city that makes it just.

This basic city is disparaged in much of the scholarly literature. Julia Annas, for example, speaks about how this city is “allowed to develop, and becomes corrupt,”<sup>10</sup> and claims that this city is fatally flawed because it is “built up purely on the basis of self-interest.”<sup>11</sup> Annas believes that Plato, in fact, reveals his lack of interest in this city by having Socrates stress that he is looking for the origins of justice and *injustice* in this and the “swollen” city which follows it, and comments,

This suggests that both the first, ‘true’ city and the corrupted city are put forward for us as models of human nature in association which display both justice and injustice, to be contrasted with the ideally just city, which displays only justice.<sup>12</sup>

Annas concludes that, in fact, “Plato has not given the first city a clear place in the *Republic*’s moral argument.”<sup>13</sup> C. D. C. Reeve sarcastically dismisses the first city as “the *Kallipolis* for money-lovers,” and says that Plato did not mean for us to take it seriously because “it includes nothing to counteract the destabilizing effects of unnecessary appetites and the *pleonexia* to which they give rise.”<sup>14</sup>

Such assessments, and the arguments offered for them, however, seem to me to be quite mistaken, and in any case clearly do not explain why Plato would represent Socrates as having such a favorable view of this city, which is never qualified or withdrawn even in retrospect in any of the remaining books of the *Republic*. But let us look more closely at the arguments offered for such negative appraisals by Plato’s commentators. Is the first city built entirely upon self-interest, as Annas says? Annas cites 369c6-7 for this claim, where Socrates says that cities come into being because people share by giving and taking, thinking that it is better for them to do so. It is significant that Socrates makes this claim, however, *before* beginning the process of “creating a city, from the beginning, in words” (369c9-10). Accordingly, this characterization of how cities come into being is not offered as a description of *only* how the *first* city comes into being, or as a subtle signal of what is *wrong* with the city that comes into being with this *raison d’être*. Rather, there is every reason to believe that this same basic motivation lies behind the generation of any city whatever, including the *Kallipolis* itself. After all, it is an unmistakable feature of Plato’s argument—and a central feature of his answer to the challenges of Glaucon and Adeimantus at the beginning of Book II—that it is in one’s interest to be just. If the first city is brought into being out of self-interest, then, and can be supposed to satisfy that

self-interest, this is an argument in favor of the first city's justice, and not against it.

Reeve, however, argues that the flaw is not the self-interest of the city's inhabitants, as such, but the fact that it is purely *material* self-interest, which leaves the city wholly unprotected against "the destabilizing effects of unnecessary appetites and the *pleonexia* to which they give rise." But Reeve imagines this without any help from Plato's own explicit description of this state, in which the relevant appetites and their satisfactions are described in ways that highlight their moderation and simplicity:

First, then, let's see what sort of life our citizens will lead when they've been provided for in the way we have been describing. They'll produce bread, wine, clothes, and shoes, won't they? They'll build houses, work naked and barefoot in the summer, and wear adequate clothing and shoes in the winter. For food, they'll knead and cook the flour and meal they've made from wheat and barley. They'll put their honest cakes and loaves on reeds or clean leaves, and reclining on beds strewn with yew and myrtle, they'll feast with their children, drink their wine, and, crowned with wreaths, hymn the gods. They'll enjoy sex with one another but bear no more children than their resources allow, lest they fall into either poverty or war. (372a5-c1)<sup>15</sup>

The citizens of Plato's first state will eschew even the luxurious excess of wearing shoes or clothing in the summers, and are careful to do nothing that would strain their very modest resources, and so will avoid both poverty (which Plato wishes to avoid even in the *Kallipolis*, as morally corrosive—see 421d13-422a3) and war. If Plato wished to portray this city as in danger of "the destabilizing effects of unnecessary appetites and the *pleonexia* to which they give rise," he has done a remarkably poor job of it!

Annas's final argument—that Socrates' urging us to see the origins of both justice and injustice in this state is a subtle sign of his lack of regard for this state—is equally flawed. First, let us recall that Socrates stated that this—the identification of the sources of justice and injustice in a city—was the goal he was seeking in bringing the city into the discussion in the first place (at 369a5-7). Second, from the fact that Socrates and his interlocutors are in a position to see what injustice in a state would be, it does not follow that the state they have envisioned is one that is filled, to any significant degree, with injustice. After all, it is precisely because Socrates and his interlocutors reach the position where

they can see the origins of both justice and injustice that they are able to conclude the discussion of the *Kallipolis* at 434b9-10: justice in the state is where each element in the state does its own, and injustice is where the elements meddle and interfere in each other's proper functioning. Third, Annas is wrong to say that, unlike the first state, the *Kallipolis* "displays only justice." If this were true, there would be no need for law courts in the *Kallipolis*, and no need for the philosopher-rulers to serve as judges for disputes that may arise among the citizens of the *Kallipolis*; but Plato tells us that there will be such needs, at 433e3-8.

What is lost, in views such as the ones we find Annas and Reeve giving, is the recognition that *none* of the states that Plato's Socrates and his interlocutors conceive of are Forms. Stated so bluntly, this may seem just obvious, but Annas' claim that the *Kallipolis* will display only justice and no injustice is one that Plato's metaphysics could never support. Only Forms are what they are and never can or do seem to be the opposite at the same time (see 478e7-479b8). Accordingly, we should expect that *both* the first city *and* the *Kallipolis* will "display" both justice and injustice, for both are only participants in, and not identical to, the Form of Justice.<sup>16</sup>

This recognition, however, shows us a further thing about how Plato's metaphysics must inform his conception of the analogy of soul and state. Insofar as Justice Itself—the Form—is one thing, then we should expect every one of its images (both in states and in individual people) to make reference to the same thing. The analogy of justice in the soul and justice in the state is secured, then, in virtue of the fact that the justice in each will resemble one another, and they will resemble one another precisely because *both* sorts of justice are images of Justice Itself. But it follows from this that *any* image of Justice will resemble any other image of Justice, in the relevant way. But states come in an indefinite number of varieties—not all are divided into three parts. Some states will be more just than others; some will have little justice in them to observe at all. Any state with detectable justice in it, however, should serve to the same degree as its justice is detectable, for Socrates and his interlocutors—and for Plato's readers—to be able to "read the large letters" of justice in the state, in preparation for an attempt to discern the same message of justice written in smaller letters within an individual. Plato's *Republic*, I am claiming, actually provides two distinct images of justice in states: in the first and most basic city, and also in the *Kallipolis*.

Of course, as Plato tells us later—when he describes how images can be used intelligently (in Books VI and VII), as they are in arithmetic and

geometry—some images are more useful as propaedeutics and heuristics than are others. In seeking to achieve a grasp of what justice is, before turning to the actual challenge itself, which was to show that justice is always preferable to injustice, Plato's Socrates must choose the most useful images of justice. When Glaucon and Adeimantus seem so scornfully unimpressed by the first city, Socrates acknowledges that a larger and more luxurious city might provide a better view of what they are looking for:

Alright, I said. I understand. It is not just how a city comes into being that we are investigating, but rather how a luxurious city comes into being. Perhaps that's not a bad thing, either. For by investigating that, we might discern how justice and injustice grow up in a city. (372e2-6)

Even if the *Kallipolis* provides a better image of Justice for the purposes of their investigation, however, as it plainly does, we should not necessarily suppose that the first city was either unjust or in any absolute way a poor or deeply flawed image of justice itself. An image may be poorly suited for an inquiry into the nature of that which the image imitates, for reasons other than that it is a poor image. Notice that in seeking to discern what justice is, Plato's Socrates insists on *not* looking first at individual human beings, but at states, for fear that the justice they sought to grasp would be too difficult to discern in individuals, without the help of the analogy to states. It plainly does not follow from this, however, that no individual human beings could ever be good or reasonably accurate images of justice.

At any rate, once justice is finally located and observed in the *Kallipolis*, Plato has Socrates repeatedly remind us that it is nothing other than the very same principle that was used in founding the first, most basic city, as well as each subsequent development worked upon that city, in transforming it into the *Kallipolis*: that each person in the city should do that and only that task in the city, for which his or her nature best suits them (see 370a7-b2, 374a4-e8, 432d7-433a6, 443b7-c7). Accordingly, it turns out that there *was* justice in the first and most basic city, which now presumably even Glaucon and Adeimantus are in a position to see: in that city, no less than in the *Kallipolis*, each person did that for which he or she was best suited by nature, and did not meddle in what others were better suited to do. Despite proving clearly (in retrospect) to be a good image of justice after all, however, the first city does not do so in virtue of any obvious division into three parts. Nonetheless, even if we granted Plato's arguments for the partitioning of the soul, we

can see now what the analogy of soul and state would consist in, if we compared Plato's tripartite soul to his first and most basic state: justice in each would be in each element or part "doing its own" and not meddling in the proper business of any other element or part. We do not have to presuppose the 3-3 specification of the soul-state analogy in order to preserve the identity of justice, and thus, the essential basis of the analogy.

### III. Justice in the Soul

If the argument of the last section is correct, Plato actually offers more than one model of the city which satisfies the analogy of justice in states and souls. Plato does not, however, provide any clear alternatives to the conception of the soul that he offers in Book IV, and so it might seem that he is as deeply wedded to his conception of the soul as tripartite as he could be to any theory he offers. But in several places in the *Republic*—including immediately before and after giving us his partitioning arguments—Plato reveals a certain lack of complete commitment to the consequences of the arguments he gives for the tripartite soul.

Socrates begins his argument for the partitioning of the soul with a word of caution:

Well, is it only a small question we have come to, O amazing one, I said, whether the soul has these three forms in it, or not?

[Glaucon responds] I do not believe it is at all a small question. For perhaps, O Socrates, the saying is true that "fine things are difficult."

Apparently, I said, but know well, O Glaucon, that in my opinion, we'll never comprehend the matter precisely using the methods we are using in our present arguments, though there is a longer and more complete way which would lead to this. But perhaps we can proceed in a way that is comparable in value to our former claims and inquiries.

By having Socrates remind Glaucon of the shortcomings of "the methods we are using in our present arguments," and noting that the same shortcomings can be expected in the arguments that follow, Plato cautions his readers that his arguments must be taken with a grain of salt.<sup>17</sup> Moreover, immediately after he gives his partitioning arguments, when he is summarizing his account of what justice is in the soul, in Book IV, Plato explicitly opens up the possibility that the soul may actually be composed of more than three parts:

The truth, however, is something more of this sort, it seems, justice is not doing one's own with regard to what is external, but with regard to what is within him, what is truly himself and his own, not allowing any part of his soul to do another's in himself, or to meddle with one another, but to arrange well what is really his own and to rule himself. He puts himself in order and befriends himself, and harmonizes the three things in him like the three harmonic limits, the low, the high, and the middle, *and any others there may be between them.* (443c9-d7)

Later, in Book IX, just when we might expect Plato's Socrates to remind us of the tripartite nature of the soul, we find him arguing, instead, that we must not suppose "that in its truest nature, the soul is this way—as embroidered in many colors and dissimilarities, and at odds with itself" (611bl-3). Socrates goes on to say that the condition in which he and Glaucon have been examining the soul is a condition much like that of the sea-god Glaucus, damaged and encrusted, as it were, in its association with the body (611b9-d8), whereas only if we could observe it in its true nature could we really discern "whether it has many parts or just one, and how and in what way it truly is" (612a3-5). Only a few pages before comparing the embodied soul to Glaucus (at 603d6-7), Socrates breezily acknowledges that the soul suffers from "countless oppositions" of the sort he took (in Book IV) to divide the soul. One wonders if all of these oppositions could be accounted for (applying *Oe*) by an appeal to only the three parts of the soul he identified in Book IV. In regard to his earlier arguments, Socrates allows only that "we have given a fitting (*ἐπιεικῶς*) account" of the soul, as it is in a human life (612a5-6).

Here, too, then, we find Plato doing the opposite of what we would expect of him if he were committed to the 3-3 specification of the soul-state analogy, for now he appears prepared to allow that the soul may not—in its ideal condition, at any rate—be composed of three parts, after all. Moreover, Socrates allows that if only they were able to observe the soul in its truest nature, they might then be able to "see justice and injustice and the other things we have been discussing more clearly" (611c4-5).

W. K. C. Guthrie has argued that Plato believed that in its truest form, the soul was not a composite after all, but, instead, a simple.<sup>18</sup> If so, because Plato claims that in its truest form, the soul would be a "much finer thing . . . than we now suppose" (611c4-5), we can infer that it would be an even better approximation or image of justice as a simple than it is in its fragmented aspect, "maimed" in its association with the body (611b10-c1).

In carefully distancing himself in this way from his earlier partitioning arguments, Plato's Socrates shows that we should not understand those earlier arguments as being in any way decisive as to how we are to conceive of "Plato's theory of the soul," if he has such a theory. Instead, we should remind ourselves what the purpose of the earlier arguments was: to try to discern the aspect of justice in a soul. In dividing the soul into three parts, Socrates and Glaucon were better able to see what it was that made a soul just, by analogy to what it was that made a state just. As such, the just tripartite soul served the discussion well, as an image of justice. But again, it does not follow from this that the image we have been given is the best possible image of justice, either in the soul or in the state. All that follows is that they are good images of justice—good enough to serve the propaedeutic and heuristic purposes for which they have been selected.

Given Plato's characterization of justice in the soul, however—each part doing its own and not meddling in the functioning of the other parts—we might well wonder how Plato could account for justice in a nonpartitioned soul, since plainly in such a case there could be no parts to "do their own." If we return to the formula that Plato has relied upon all along, however, we will see that this worry is no real threat to his conception. Justice in a composite thing—either soul or state—is all along conceived as nothing other than the overall harmony that is achieved in the composite in which each element of the composite "does its own." States are just, Plato tells us, when and insofar as each element within the state "does its own." In the first and most basic state, and also in the *Kallipolis*, this ultimately amounts to the situation in which each individual does what he or she is naturally best suited to do, and does not meddle in others' affairs. Any single individual in the state, therefore, will be civically just insofar as he or she "does his or her own." After all, it is not as if the state can somehow be just in which the citizens themselves act unjustly. Similarly, just as we might talk about certain impulses and motives being unjust and others being just, Plato can say that each part of a soul is just when and only when that part is doing its own, and he can even say that individual desires and motives are just when and only when they are symptomatic of the soul or its relevant part "doing its own." Since the justice of a composite consists in the harmony resulting from the justice of each of its component (including, presumably, its indivisible) parts "doing its own," it follows that indivisible simples are just when and to the degree that they "do their own."

But the "doing one's own" formula ordinarily applies to activities

that are, as the Greeks would say,  $\pi\rho\circ\varsigma\tau\iota$ , or “in relation to something.” For a carpenter in Plato’s *Kallipolis* to “do his own” is not just for him to do carpentry, but also for him not to meddle in others’ work. For one’s “spirited part” to be just, it must not only provide the right motivations, it must also not interfere in one’s reasoning or appetites. A single individual stranded on a desert island, accordingly, would not have the opportunity to engage in the civic form of justice (even if she might still be psychically just), since she would have no “other” relative to whom she could “do her own.” Similarly, it might be supposed, if we conceive the soul as a simple, we lose any opportunity to characterize it as just, because there are no longer any others (in this case, parts of the soul) for the soul to be just in relation to.

But Socrates gives an argument in Book I that neutralizes this concern, proclaiming that the soul has a function, and that the virtue of a soul would consist in the soul performing its function well. The function of the soul is “management, rule, deliberation, and all other such things” (353d4-6), and to this list Socrates immediately adds “living” (at 353d9). Justice is the virtue of the soul, and so the just soul will be one that performs these functions well (353d11-e11). The “doing one’s own” formula, accordingly, may refer to the ways in which one functions in regard to others, but it can also, we must suppose, apply to the proper functioning of the soul in regard to the soul’s most basic functions, which include most importantly the management of one’s own life.

#### IV. Concluding Remarks

In the first part of this paper, I claimed that Plato’s arguments for the partitioning of the soul are unreconstructable failures. If the analogy of justice in the state and in the soul required what I have called the 3-3 specification of that analogy, this failure would infect and ruin much of the argument of the *Republic*. In fact, however, I have argued that it does not do so, and I gave reasons, in the second and third parts of the paper, respectively, why even Plato himself was not committed in any deep way (in the *Republic*, at least) either to the necessity of a tripartite state, for the state to be just, or to the theory of the tripartite soul, or for this conception of the soul to be a requirement for the soul to be just. I have not claimed that Plato remained noncommittal about his tripartite analysis of the soul, which shows up prominently again in the *Phaedrus* and the *Timaeus*, though even in these dialogues, I believe his commitment to the

analysis may be doubted. An argument for this claim, however, must wait for another day. Even if Plato did come to accept his partitioning arguments without reservation, it would not count as evidence against my argument in this paper, for I have claimed only that we should not accept them, that Plato does not actually require such arguments to succeed in order to defend his soul-state analogy, and that various things he says show that Plato himself was aware that the partitioning of the state and soul were not required by the soul-state analogy in the *Republic*.

At base, my argument is founded upon my claim that Plato’s metaphysics required that there be a single, unitary, and unequivocal conception of justice, which Plato would tie to his idea that there was a single pure Form of Justice. Even this part of my argument is not uncontroversial, however, for scholars have often talked of there being two conceptions of justice in Plato’s *Republic*. The most troubling case of this attribution was made in the famous paper by David Sachs,<sup>19</sup> in which Sachs argued that Plato’s defense of justice rests on a fallacious equivocation between the two senses of justice, which Sachs called the “vulgar” and the “Platonic” conceptions. But even those who have sought to defend Plato against Sachs<sup>20</sup> have allowed that there are two distinct conceptions of justice—usually identified as the social or civic conception, in which justice is conceived in terms of the qualities of the performance or abstention from certain sorts of actions, and the psychic conception, in which justice is conceived in terms of balance or harmony in the soul. Plato’s defenders have generally, and I think rightly, argued that Plato could show that those who were psychically just would also be socially or civically just. Although I do not deny that individuals whose souls are just will act justly—after all, Plato explicitly says this at 443c9-444a2—I do think it is a mistake to say that there is more than one conception of justice ever at work in Plato’s *Republic*.

Justice, for Plato, is ultimately nothing other than the Form of Justice. Any instantiations of this Form, and images of it, will be just if, and only to the degree that, they approximate or participate in the character of this Form. But the character of this Form does not change in relation to its different sorts of participants: whatever their other differences may be, insofar as these participants are just, they will be just in the same respect, as it were. Anything that approximates or images the Form of Justice will resemble anything else that does, and so may be found to be analogous to anything else that approximates or images the Form of Justice, but the resemblance will always be on the basis of the same Form and therefore in virtue of whatever it is in the two approximations or im-

ages that displays the relevant character associated with this single Form. Accordingly, my argument has been that Plato does not need to be committed to any specific theory of the soul or of the state, to make an analogy between the instantiations of justice that may be found within each.

In each and every instance of justice represented to us in the *Republic*, we are told that the justice we are supposed to observe derives from and is present in virtue of that instance's proper functioning. It is the proper functioning of the soul that we are told is justice in the soul, in Book I. It is the proper functioning of each of the citizens in the first and most basic state that makes these citizens and their state just, in Book II. It is the proper functioning of each of the citizens in the much expanded state, the *Kallipolis*, and each of the classes of this state, that we are told make the citizens and classes of this state—and the whole state itself—just, in Book IV. Also in Book IV, we are told that it is the proper functioning of the parts of the soul that make these parts and the activities they motivate just, and which makes the whole soul and the person just. In every case, justice as proper functioning can be understood under the formula “doing well what the thing is naturally suited to do, and not attempting to do what only some other thing is better suited to do.” But in all of its instantiations, justice is proper functioning, for Plato. There may appear to be many distinct characterizations of this, just as Plato recognizes many images of the single Form. There is a unity that runs through all of the characterizations, however, and the metaphysical basis for this unity is none other than the Form.<sup>21</sup>

## Notes

1. For examples, according to David Sachs, “A Fallacy in Plato’s *Republic*,” *Philosophical Review* 72 (1963): 141-58 and Gregory Vlastos, “Justice and Happiness in the *Republic*,” in *Platonic Studies*, 2d ed., ed. Gregory Vlastos (Princeton, N.J.: Princeton University Press, 1981), 111-39, Plato equivocates on the conception of justice, though they offer different accounts of the equivocation and come to different conclusions about how serious such equivocations are to Plato’s overall argument.

2. Other criticisms of Plato’s psychology in the *Republic* may be found in the literature, as well. For example, Christopher Bobonich, “Akrasia and Agency in Plato’s *Laws* and *Republic*,” *Archiv für Geschichte der Philosophie* 76 (1994): 3-36 and Terence Irwin, *Plato’s Ethics* (Oxford: Oxford University Press, 1995), 217-22, argue that Plato’s conception of the parts of the soul mis-

takenly attributes to them a kind of agency which belongs properly only to the individual of which they are parts. In this paper, I do not address such concerns directly, but I would note that insofar as my arguments give some reason to think that Plato was not wholly committed to his partitioning arguments, he cannot either be regarded as wholly committed to the sorts of errors Bobonich and Irwin have identified. Yet other problems regarding Plato’s psychological account (and thus its application to the soul-state analogy) are noted in Mitchell Miller, “Platonic Provocations: Reflections on the Soul and the Good in the *Republic*,” in *Platonic Investigations*, ed. D. J. O’Meara (Washington, D.C.: The Catholic University Press, 1985), 163-93, and Lloyd Gerson, “A Note on Tripartition and Immortality in Plato,” *Apeiron* 21 (1987): 81-96. I regret that I cannot directly address or incorporate all of their arguments or the issues they raise in this paper.

3. I am indebted, in much of the analysis that follows, to suggestions by Fred D. Miller, Jr., and to subsequent discussions with a former student, Michael Fiebig, whose noble and thoughtful persistence in trying to defend Plato against my criticisms helped me to sharpen those criticisms considerably.

4. For this reading of  $\chi\alpha\tau\alpha\tau\alpha\tau\alpha\tau\alpha$  at 436b8, see R. F. Stalley, “Plato’s Argument for the Division of the Reasoning and Appetitive Elements within the Soul,” *Phronesis* 20 (1975): 110-28.

5. Fred D. Miller, Jr., argues for this interpretation in Fred D. Miller, Jr., “Plato on the Parts of the Soul,” in *Plato and Platonism*, ed. Jan Max van Ophuijsen (Washington, D.C.: The Catholic University of America Press, forthcoming).

6. Anthony Kenny appears to understand Plato’s partitioning principle in this way, as “the principle of non-contrariety” (Anthony J. Kenny, “Mental Health in Plato’s *Republic*,” *Proceedings of the British Academy* 55 [1964]: 229-53).

7. Terry Penner, “Thought and Desire in Plato,” in *Plato II: Ethics, Politics, Philosophy of Art and Religion*, ed. Gregory Vlastos (New York: Anchor Books, 1971), 108-11, 113-16 and J. F. Tiles, “The Combat of Passion and Reason,” *Philosophy* 52 (1977): 322 discuss this case, with differing conclusions. Tiles appears to assume that Plato’s partitioning principle is what I have called *Oe*.

8. I have only given an example that would partition Plato’s appetitive part of the soul, if we apply *Oe*. But I believe similar examples could be given to partition the other parts, as well. For example, consider a case where one finds oneself attracted to the honor one could win in some competition, but honorably repelled by the fact that the only hope of winning one had would be by using banned performance-enhancing drugs. Or, one finds oneself attracted to a certain philosophical doctrine’s explanatory potential, while also feeling doubts about its foundational assumptions. Such cases, I suggest, would partition the high spirit and the rational parts of the soul, respectively, if we applied *Oe*.

9. See Daryl H. Rice, *A Guide to Plato’s Republic* (Oxford: Oxford University Press, 1998), 58: “Socrates argues that if his analogy between the city and

the soul of the individual is to be helpful, cities and individuals must be morphologically similar; that is, they must have the same basic structure, or form."

10. Julia Annas, *An introduction to Plato's Republic* (Oxford: Oxford University Press, 1981), 76.

11. Annas, *Introduction*, 78.

12. Annas, *Introduction*, 78.

13. Annas, *Introduction*, 78.

14. C. D. C. Reeve, *Philosopher-Kings: The Argument of Plato's Republic* (Princeton, N.J.: Princeton University Press, 1988), 171.

15. To highlight the discrepancy between Reeve's assessment of the first city, and Plato's actual description of it, I have used Reeve's own translation here (G. M. A. Grube and C. D. C. Reeve, *Plato: Republic* [Indianapolis, Ind., and Cambridge, Mass.: Hackett, 1992], 47).

16. This also, presumably, accounts for what would otherwise be the very strange qualification Socrates makes, at 444a6, that in identifying the just man and the just city as they have, they have not told a "complete falsehood." There would be no falsehood at all, in what they have said, if they had identified the Form of Justice, rather than merely good images of it.

17. I owe this observation to Mitchell Miller, who makes a similar point in Mitchell Miller, "Platonic Provocations," 169.

18. W. K. C. Guthrie, "Plato's View of the Nature and the Soul," in *Plato II: Ethics, Politics, Philosophy of Art and Religion*, ed. Gregory Vlastos (New York: Anchor Books, 1971), 230-43.

19. Sachs, "A Fallacy in Plato's *Republic*."

20. See, for examples, R. Demos, "A Fallacy in Plato's *Republic*?" *Philosophical Review* 73 (1964): 345-98; L. Galis, "The State-Soul Analogy in Plato's Argument that Justice Pays," *Journal of the History of Philosophy* 12 (1974): 285-93; R. Hall, "Plato's Political Analogy: Fallacy or Analogy?," *Journal of the History of Philosophy* 12 (1974): 419-35; R. Sartorius, "Fallacy and Political Radicalism in Plato's *Republic*," *Canadian Journal of Philosophy* 4 (1974): 349-63; J. Schiller, "Just Man and Just Acts in Plato's *Republic*," *Journal of the History of Philosophy* 6 (1968): 1-14; Vlastos, "Justice and Happiness in the *Republic*"; S. Waterlow, "The Good of Other in Plato's *Republic*," *Proceedings of the Aristotelian Society* 47 (1973): 19-36; and R. Weingertner, "Vulgar Justice and Just Acts in Plato's *Republic*," *Philosophy and Phenomenological Research* 25 (1964/65): 248-52. A few scholars (in my view rightly) do not make such a concession—see, for example, Julia Annas, "Plato and Common Morality," *The Classical Quarterly* 28 (1978): 437-51; Scot Yoder, "Justice in Plato's *Republic*: Constructive Moral Philosophy vs. Moral Apologetics," Society for Ancient Greek Philosophy, Binghamton, N.Y., October, 1995.

21. I presented various earlier versions of this paper as one of the Patricia Crawford lectures at San Diego State University (September, 1998), at the Third Arizona Colloquium in Ancient Philosophy (February, 1998), and as colloquia presentations at the University of Toronto (December, 1997) and at Michigan

State University (September, 1998). I am indebted to my hosts in San Diego (especially Angelo Corlett and Mark Wheeler) and to those in Toronto (especially Brad Inwood and Lloyd Gerson) for their searching questions and criticisms. I gratefully acknowledge, as well, the assistance I have received from Mitchell Miller, my commentator at the Third Arizona Colloquium in Ancient Philosophy, who provided in private correspondence several detailed suggestions, in addition to those he read at this conference, several of which I have incorporated into the current version of the paper. I am also grateful to my colleagues at Michigan State University, and to the members of my Fall 1997 seminar there—Timothy O'Neill in particular—for their many comments and suggestions. I have also benefited from long discussions of this topic with Fred D. Miller, Jr. and with Michael Fiebig, both of whom were especially helpful in leading me to sharpen my criticisms of Plato's partitioning arguments. Finally, I acknowledge the help I have received from Rachana Kamtekar and Ellen Wagner, who read and commented extensively on the version of the paper I gave in Arizona. None of the above can be assumed to agree with my arguments, and all remaining errors in the paper are mine alone.

## Simple Souls

Christopher Shields

### I. The Question of Soul Division in Plato's Development

According to a credible account of Plato's development, *Republic* iv contains an argument intended to overturn an idiosyncratically Socratic thesis concerning the possibility of *akrasia*. Socrates flatly denies that *akrasia* is possible. This he does most directly in the *Protagoras*, where he asserts that "those things which one regards as bad, one neither goes toward nor accepts willingly" (ἀ δὲ ἡγεῖται κακά, οὐδένα οὔτε ἵέται ἐπὶ ταῦτα οὔτε λαμβάνειν ἔκόντα; 358e). His reasons for advancing this initially counterintuitive claim are varied, but evidently include a narrowly cognitive conception of human motivation coupled with a broadly egoistic form of psychological hedonism. These claims jointly result in the thesis that anytime I choose to do *x* over *y*, when doing *y* offers me greater overall pleasure than doing *x*, I act on the basis of a mistaken calculation. Having recognized doing *x* as possible for me, I misreckon its net pleasure relative to the pleasure to be had by my doing *y*. My choice to do *x* is thus due to a simple cognitive mistake, a mistake in calculation akin to the mistake I make when I judge that the area of a square half the area of square *ABCD* is a square whose legs are half as long as those of *ABCD*. When corrected, I see my mistake and I adjust my attitudes and behavior accordingly.

On this developmental account, key to Plato's rejection of Socrates' point of view is his recognition of a kind of complexity internal to the

soul which Socrates had failed to appreciate, or at any rate had failed to acknowledge.<sup>1</sup> This seems to be the point of the admonition Plato issues just after establishing the first two parts of soul in *Republic* iv:

Μήτοι τις, ήν δ' ἐγώ, ἀσκέπτους ἡμᾶς ὄντας θορπυβήση, ὡς οὐδεὶς ποτοῦ ἐπιθυμεῖ ἀλλὰ χρηστοῦ ποτοῦ, καὶ οὐ σίτου ἀλλὰ χρηστοῦ σίτου. πάντες γάρ ἄρα τῶν ἀγαθῶν ἐπιθυμοῦσιν· εἰ οὖν ἡ δίψα ἐπιθυμία ἐστί, χρηστοῦ ἀν εἴη εἴτε πώματος εἴτε ἄλλου ὅτου ἐστὶν ἐπιθυμία, καὶ αἱ ἀλλαι οὕτω.

Let no-one, I said, unsettle us when we are unwary, by claiming that nobody desires drink, but rather good drink, nor food, but rather good food, on the grounds that after all everyone desires good things. If, then, thirst is a desire, then it would be a desire for good drink or whatever else it is a desire for, and so on for the other desires (*Rep.* 438a).

This sounds, at first pass, like a straightforward repudiation of the Socratic suggestion that all apparent cases of *akrasia* must be resolved to be cases of poor arithmetic. Some of my desires, those which well up in my *epithumētikon*, are *good-indifferent*. These desires can actually conflict with the deliverances of my rational faculty, my *logistikon*. If I act upon them, and against my best interest, the explanation will not typically be cast in terms of a mere miscalculation. Instead, my reason will have been weak, and will have yielded to a more powerful soul part, my faculty of desire. This, then, affirms the very thesis Socrates sought to reject in the *Protagoras*. There he had sought to deny the view of the many, according to which “though knowledge is often present in a man, it does not rule over him, but rather something else does, sometimes desire, sometimes pleasure, sometimes pain, other times love, often fear” (ἐνούσης πολλάκις ἀνθρώπῳ ἐπιστήμης οὐ τὴν ἐπιστήμην αὐτοῦ ἀρχειν ἀλλ' ἄλλο τι, τοτὲ μὲν θυμόν, τοτὲ δὲ ἡδονήν, τοτὲ δὲ λύπην, ἐνίοτε δὲ ἔρωτα, πολλάκις δὲ φόβον; 352b-c).

This developmental story has two tiers, one at the level of Plato’s moral psychology and the other, deeper, at the level of his metaphysical psychology. The two tiers are related: precisely by recognizing complexity internal to the soul itself, Plato is able to allow for the possibility of *akrasia*. Indeed, the thesis about the soul’s nature both makes possible and renders explicable a phenomenon Socrates had sought to deny and explain away. So, the moral psychology and the metaphysical psychol-

ogy march in step. A simple soul gives way to a complex soul; and a Socratic paradox is dissolved.

This two-tiered account thus embraces three distinct theses, one psychological, one moral-psychological, and one regarding the relation between the first two. The first thesis holds:

(I) The early Plato, especially the Plato of the *Phaedo*, is like Socrates before him in supposing that the soul is simple; he later abandons this view, most forcefully in *Republic* iv, in favor of the view that the soul is composite.

The second maintains:

(II) Plato, unlike Socrates, accepts the fact of *akrasia*.

And, finally, the third connects these two claims:

(III) (I) explains (II). That is, Plato accepts, and indeed *can* only accept, the fact of *akrasia* because he comes to reject the simplicity of the soul in favor of a composite view.

Taken together these three theses constitute what I will call the standard developmental account.

The standard developmental account, I believe, is incorrect. Indeed, it is multiply problematic, because every single one of the theses which constitute it is dubious. To begin, the thought that Plato’s introduction of a divided soul underwrites a noteworthy development in his moral psychology strikes me as problematic, both in its own terms and as an account of Plato’s development. In the first instance, a divided soul is neither necessary nor sufficient for the possibility of *akrasia*. That is, one could accept a bi- or tripartite psychology without countenancing the possibility of *akrasia*. I might, for example, be a psychological hedonist who holds, as a matter of psychological fact, that everyone always maximizes his or her own perceived pleasure. That I recognize various divisions within the soul would then be neither here nor there with respect to the possibility of *akrasia*. Conversely, I might hold to a simple, unified soul, but nevertheless insist that *akrasia* is fully possible. On this scenario, I would simply maintain that though I judge that, all things considered, *x* is better than *y*, I nevertheless find myself doing *y* because of the allure *y* holds for me. I later upbraid myself in the normal way; but

I never fault one part of myself while promising another to do better next time by letting that other part rule.<sup>2</sup> Given that soul-complexity and *akrasia* can be thus decoupled, we should not move too readily to the conclusion that Plato came to acknowledge the possibility of *akrasia* because he came to recognize complexity in the soul. That is, we should already be skeptical about (III) on purely conceptual grounds.

Moreover, there are nonconceptual reasons for doubting this story as an account of Plato's development in psychology, because there are textual impediments to (II), the thesis that Plato accepts but Socrates rejects the possibility of *akrasia*. First, (II) assumes that the Socrates of the *Protagoras* rejects the possibility of *akrasia* in *propria persona*; second, it presupposes that the Plato of *Republic* iv rejects the Socratic position on *akrasia*. But both of these can be assailed in various ways. On the *Republic* side, it is noteworthy that not long before offering the analogy of the sun, Plato says, rather blithely:

ὅ δὴ διώκει μὲν ἀπασα ψυχὴ καὶ τούτου ἔνεκα πάντα πράτει, ἀπομαντευομένη τι εἶναι, ἀποροῦσα δὲ καὶ οὐκ ἔχουσα λαβεῖν ἵκανως τι ποτ' ἔστιν οὐδὲ πίστει χρήσασθαι μονίμῳ οἴᾳ καὶ περὶ τᾶλλα, διὰ τοῦτο δὲ ἀποτυγχάνει καὶ τῶν ἔλλων εἴ τι ὅφελος ἦν

[The good is] that which every soul pursues and on account of which it does all things. Divining that the good is something, but being perplexed and incapable of grasping sufficiently what it is or of having the steadfast confidence it has concerning other things, it misses the benefit, if there was one, from those other things (505e-506a).

On its surface, this passage seems easily compatible with the rejection of *akrasia* of *Protagoras* 352b-358e, a passage thought to be the canonical expression of the Socratic paradox. Indeed, it seems more than compatible: if in fact every soul does everything it does on account of the good, then it seems that it fails to meet its end only because of a cognitive confusion, because it is, as Plato here says, perplexed. So, there is reason to doubt (II), and hence already additional reason to doubt the explanatory hypothesis offered in (III).

The same holds true for (I), a thesis which ascribes to the *Protagoras* a view more readily found in the *Phaedo*. That is, this account assumes that the *psuchē* in the *Protagoras* is simple, a thesis not defended or even articulated in that work. So here too it is worth investigating the evidence for ascribing to Socrates a view according to which the soul is non-

composite; if we find no such evidence, we will have still further reason to doubt (III). For absent (I) and (II), (III) advances an explanation where none is wanted.

However that may be, it may nevertheless seem safe to understand (I) in its own terms, whether or not it is understood as a plank in an otherwise problematic developmental account. For surely Plato accepts a simple soul in the *Phaedo* only to reject any such view, in the strongest possible terms in *Republic* iv. A simple soul is incompatible with the complexity exhibited in the argument of *Republic* iv. Here at any rate, we have a clear change of mind.

I wish to argue that insofar as Plato accepts a simple soul in the *Phaedo*, he continues to do so in the *Republic*. I want to argue more narrowly, that is, that the “soul-division” of *Republic* iv is fully compatible with the soul's simplicity, and that, consequently, (I) is incorrect. If that is so, there is still less reason to posit (III). In fact, Plato's argument for a tripartite psychology at *Republic* iv 436b-441a *does not and could not* generate a soul with essentially distinct parts. Accordingly, his argument for a tripartite psychology in *Republic* iv cannot be understood to advocate a conception of the soul in any way incompatible with his characterizations of the soul as simple given elsewhere in the corpus, most strikingly in the *Phaedo* and in *Republic* x.<sup>3</sup>

## II. Psychic Simplicity and Immortality in the *Phaedo* and *Republic* x

In the Affinity Argument of the *Phaedo* (78b-84b), Socrates links indestructibility and incompositeness, especially at 78c1-4 and 80b2. Because the soul is akin to the Forms in being “most similar to what is divine, immortal, intelligible, uniform, indissoluble, and always similarly disposed with respect to itself” (τῷ μὲν θείῳ καὶ ἀθανάτῳ καὶ νοητῷ καὶ μονοειδεῖ καὶ ἀδιαλύτῳ καὶ ἀεὶ ὠσαύτως κατὰ ταῦτα ἔχοντι ἔαυτῷ ὅμοιότατον εἶναι ψυχή; 80b1-2), it too must be completely indissoluble, or something close to being so (ψυχὴ δὲ αὐ τὸ παράπαν ἀδιαλύτῳ εἶναι η ἐγγύς τι τούτου; 80b10). The argument, whose modest conclusion befits its analogical terms, relies in part on the thought that only what is incomposite (ἀσύνθετον) by nature escapes the dissolution which infects the visible world (78c1-4). So, Plato argues: (i) the soul is incomposite; (ii) what is incomposite cannot suffer

dissolution; (iii) what cannot suffer dissolution is immortal; (iv) so, the soul is immortal.

Indeed, Plato relies not merely on the thought that only what perishes is composite but that everything composite is liable to dissolution. With the tacit addition of a disputable presupposition that whatever is liable to dissolution *will* at some point be dissolved, Plato is at liberty to conclude that all and only those things which are incomposite are immortal. This biconditional, then, underscores the entire Affinity Argument:

Nec. (x is incomposite iff x is eternal).<sup>4</sup>

With this biconditional in place it is easy for Plato to bifurcate the visible and the invisible as he does. All those things which are visible are extended in space, and so composite, and so perishable. By contrast, whatever is invisible is abstract, and not arrayed in space, and so incomposite, and so eternal. The soul, it is urged, is akin to the Forms in just this way.

The argument is relatively straightforward, simple, and in some ways attractive, relying as it does on a premise which many others have found tempting, that simplicity precludes destruction, since destruction requires resolution into parts.<sup>5</sup> Simple things being as such partless cannot go out of existence.

However that may be, the argument employs a conception of the soul which has struck many as manifestly incompatible with the conception at work in the *Republic*.<sup>6</sup> The *Phaedo* relies on the simplicity of the soul. *Republic* iv argues at length that the soul has three parts. Nor is this argument incidental. As Plato notes, it is crucial to the entire argument of the *Republic* itself that the soul be shown to have three parts; otherwise it could not be shown to be isomorphic with the state whose justice was introduced precisely to illuminate the nature of justice in the individual (*Rep.* 368d-369a, 434d-435b). So with the acceptance of an isomorphism between soul and state comes a repudiation of the soul's simplicity.

Plato disagrees; indeed, in *Republic* x, using language in many ways reminiscent of the *Phaedo*, he once more comes around to the view that the soul "is akin to what is divine, deathless, and always is" (611e2; συγγενής οὖσα τῷ τε θείῳ καὶ ἀθανάτῳ καὶ τῷ ἀεὶ ὄντι). More to the point, he does so precisely by way of correcting what he thinks might be an easy misconception derivable from his discussion of the soul's internal complexity earlier in the *Republic*:

μήτε γε αὖ τῇ ἀληθεστάτῃ φύσει τοιοῦτον εἶναι ψυχήν, ὅστε πολλῆς ποικιλίας καὶ ἀνομοιότητός τε καὶ διαφορᾶς γέμειν αὐτὸν πρὸς αὐτό.

πῶς λέγεις; ἔφη.

οὐ δάδιον, ἦν δ' ἐγώ, ἀδίδιον εἶναι σύνθετόν τε ἐκ πολλῶν καὶ μὴ τῇ καλλίστῃ κεχρημένον συνθέσει, ὡς νῦν ἡμῖν ἐφάνη ἡ ψυχή.

"nor <must we think> that the soul, in its truest nature, is this sort of thing, so that it is full of variation and dissimilarity and difference itself in relation to itself."

"What do you mean?" he said.

"It is not easy," I said, "for what is composite and from many things and not most finely fitted to be eternal, as the soul appeared to us to be" (611b).

Plato corrects this misconception with a striking comparison of the soul to the barnacle-encrusted sea god Glaucus, whose true nature we can scarcely ascertain, owing to its degradation and defilement by the ravages of the ocean. So the soul has been savaged by the evils attendant upon its incarnation. The account of its parts offered in the *Republic* is now held to be adequate when its embodied state is in view; when stripped of its corporeal accretions, the soul's condition, whether simple or with parts, becomes manifest (611e).

The comparison with Glaucus bespeaks a willingness on Plato's part to acknowledge an incompatibility between two conceptions of the soul, one simple and one complex. Plato does implicitly allow that if the soul were in its nature genuinely complex, that if it had genuine parts, it would in fact be corruptible. This, though, only underscores and emphasizes a continuity between the *Phaedo* and the *Republic*. For he continues to maintain the same biconditional pairing of eternality and incomposite-ness that he had relied upon in the *Phaedo*. He feels compelled to defend himself by insisting that the picture of the complex soul deployed throughout *Republic* iv-x is a picture of a soul in its sullied state. And its association with the body is here, as in the *Phaedo*, the culprit.

There is a temptation to think of the comparison to Glaucus as an unhappy expedience, as an attempt to salvage a manifestly lame argument for the soul's immortality whose obvious incongruence with the rest of the *Republic* Plato could not but fail to notice.<sup>7</sup> That is, Plato first advances the so-called Special Vice Argument for the soul's immortality in *Republic* x (608c-611a),<sup>8</sup> an argument which prompts him to wonder

how an immortal entity might yet be composite (611b), and then proceeds to treat the tri-partite psychology of the rest of the *Republic* as if it were somehow illusory, or at least inessential. When washed clean by philosophy, the true soul will be revealed simple and wisdom-loving; it will be, in short, the soul of the *Phaedo*, a soul effectively exhausted by its intellectual capacity. So, in a misguided and somewhat desperate attempt to rescue an incurably defective argument, Plato backpedals, treating the entire psychological edifice of the *Republic* as expendable.

This interpretation has one thing right: it treats the simplicity of the soul as incompatible with one form of complexity. Plato evidently himself sees just this problem; this is why he raises the concern. But the attempt simply to set aside Plato's Glaucus-image as unfortunate is unwarranted, obscuring as it does a genuine continuity in Plato's thought. To begin, Plato does not conclude in *Republic* x that the true soul is *in fact* or somehow *demonstrably* simple. Instead, he allows that when freed of its association with the body, the soul would present its true nature, where what that nature might be remains an open question.<sup>9</sup> If the soul were cleansed, he suggests, "then one might see its true nature, whether it is multiform or uniform, and whether it is such and how" (612a; *καὶ τότ' ἂν τις ἴδοι αὐτῆς τὴν ἀληθῆ φύσιν, εἴτε πολυειδῆς εἴτε μονοειδῆς, εἴτε δηπή ἔχει καὶ δπως*). This echoes the caution of the *Phaedo* almost exactly, since there, even in the Affinity Argument, Plato had not wanted to insist categorically that the soul is indissoluble, since he qualified his conclusion by saying that it must be so *or nearly so* (*Phaedo* 80b10). In both cases, we find not dogmatism from Plato, but a caution which befits the difficulty of the subject. Plato is free to rely on an argument from simplicity, relying on the analytic premises, while remaining cautiously optimistic about whether the soul qualifies by being simple in the requisite sense.

### III. The Mereology of Soul Division in *Republic* iv

These forms of continuity may prove unfortunate for Plato. If we grant that both the *Phaedo* and *Republic* x presuppose, with due precautions, that the essential soul is simple, then we find him in an awkward position. Sandwiched between these two chapters we have the bulk of the *Republic*, including most notably *Republic* iv, which contains an argument to show precisely that the soul is complex and not simple.<sup>10</sup> This argument, if sound, seems to require that the soul have parts not merely

in the way that Glaucus has encrustations, that is contingently, but rather essentially. If so, according to this approach, the attempt in *Republic* x to soft-pedal the mereology of *Republic* iv and following will be wholly ineffectual.<sup>11</sup>

I want to argue that nothing in Plato's argument for soul division is incompatible with the picture of the soul offered in *Republic* x.<sup>12</sup> One test of whether I am correct should be, then, whether the soul described in *Republic* iv qualifies as a suitable candidate for an argument from simplicity. I argue that it does. In short, the argument for psychic division does not yield parts of a sort which would be incompatible with either the Affinity Argument of the *Phaedo* or the Special Vice Argument of *Republic* x.

The argument for psychic division in *Republic* iv proceeds as follows:<sup>13</sup>

1. The same thing cannot do or undergo contraries with respect to the same element of itself (436b8). (I will call this *PGP*, or the *part generating principle*.)
2. Acceptance and pursuit are contraries of rejection and avoidance (437b1-5).
3. Instances of acceptance and pursuit are: appetite (hunger or thirst), willing, and wishing (437b7-c7).
4. Instances of rejection and avoidance are: refusal, unwillingness, and lack of appetite (437d8-10).
5. So, if we find an instance of someone both having an appetite for *x* and refusing *x*, we will have an instance of someone standing in contraries with respect to *x*.
6. In fact, our souls sometimes both have an appetite to drink and refuse to drink (439c1-2).<sup>14</sup>
7. So, we have an instance in which our souls undergo contraries with respect to *x*.
8. Since, by (1), nothing can undergo contraries with respect to the same element of itself, whenever any soul is in the condition described in (6), it must be with respect to different elements in that soul.
9. Hence, our souls have distinct elements.

In describing the conclusion in terms of distinct "elements," I intend to be neutral with respect to various comparatively determinate alternatives. If we think that the argument establishes distinct *proper parts* of the soul,

then we are already in danger of supposing that it establishes a form of complexity incompatible with simplicity. By contrast, if we think immediately that it establishes only distinct *properties*, or *aspects*, or *features*, or *types of psychic phenomena*, then we have already presumed that the argument establishes not proper parts, but only distinct kinds of qualities in the soul. It would, however, be incorrect to make any of these determinations before asking about the kinds of parts or features PGP generates. It is to this question that I now turn.

In the rich literature on the matter of soul-division in *Republic* iv, the vast majority focuses on an important question concerning (1), whether it is intended as a logical or formal principle, perhaps an instance of the principle of noncontradiction, or some kind of synthetic principle whose truth must be independently established.<sup>15</sup> There is surprisingly little which addresses the following simple question, which is for our purposes far more important: what kind of elements or parts does this argument, if sound, establish?<sup>16</sup> A weaker question might also be asked: even if unsound, what sorts of elements or parts might this argument reasonably be understood to be attempting to establish? These are questions about Plato's mereology; and they are precisely the questions we need to answer if we are to ascertain whether the soul here divided is nevertheless simple in the way requisite for an argument from simplicity.

For clarity's sake, let us distinguish the following notions of *part*. This will enable us to formulate our question with increased precision. In these definitions, I take as primitive the notion of being a *portion* of something. This notion can be given a precise definition, but that is not necessary for the present purposes. More important is the stipulation that the notion of a portion is sufficiently broad to encompass a variety of different kinds of parts, including physical components, properties, aspects, and abstract features, including limits and boundaries:

- *Aggregative part*:  $x$  is an aggregative part of  $y$  iff: (i)  $x$  is a portion of  $y$ ; (ii)  $x$  can exist as  $x$  after the dissolution of  $y$ .
- *Organic part*:  $x$  is an organic part of  $y$  iff: (i)  $x$  is a portion of  $y$ ; (ii)  $x$  is a functionally defined entity; and (iii)  $x$  is parasitic on  $y$  for its identity conditions.
- *Conceptual part*:  $x$  is a conceptual part of  $y$  iff: (i)  $x$  is a portion of  $y$ ; (ii)  $x$  is not a functionally defined entity; and (iii)  $x$  is parasitic on  $y$  for its identity conditions.

Some illustrations:

- Aggregative parts: a marble in a pile of marbles; an apple in the set of all apples; a book in Bodleian.
- Organic parts: my heart; a dollar in a monetary system; a carburetor; a door; the president of a university.
- Conceptual parts: The Louvre's top half; the center of Ireland; Trevor's baldness; Petra's degree of sympathy; the terminus of line segment  $AB$ ; Alcibiades' appearance; the perceptions of a Leibnizian monad.

These illustrations are meant to give a some indication of the kinds of parts which might be at play in Plato's argument for soul division.

With this framework in place, I offer two arguments, the first concerning the form of mereology assumed in the argument and the second concerning the question of whether the argument establishes *essential* parts of the soul, however those parts are construed. The first argument is as follows:

1. Plato's argument for soul-division in *Republic* iv establishes at most conceptual parts.
2. If (1), that argument is compatible with the form of simplicity required for the Affinity Argument of the *Phaedo* and the Special Vice Argument of *Republic* x.
3. Hence, Plato's argument for soul-division is compatible with a simple soul.

I will comment on (1) and (2) in turn.

Two points are relevant to our assessment of (1). The first concerns the fact that the first premise of Plato's argument for soul division cannot be restricted to aggregative or organic parts. However it is to be understood, as a formal or a synthetic principle, the first premise (that the same things cannot do or undergo contraries with respect to the same element of itself [436b8]), shows at most that the one entity must at times have distinct conceptual parts. Thus, for example, we might say that point  $A$  is the terminus of two lines, one curved and one straight.  $A$  itself thus has two elements, being the terminus of a straight line and being the terminus of a curved line. Since no straight line is also curved, these features in  $A$  must be distinct. But  $A$  has no aggregative or functional parts. In fact, such parts as  $A$  has are all conceptual parts. Given that Plato's principle generates just these kinds of parts, it cannot, as such, generate other forms of parts.

The second point about this first premise is that this is a consequence of his principle which Plato himself fully appreciates. When canvassing putative counterexamples to the principle of contraries, Plato considers the case of a spinning top, which might be thought to be both spinning and standing still. His response is to distinguish two elements in tops: "We would say that there is an axis and a circumference in them" (φαῖμεν ἀν ἔχειν αὐτὰ εὐθύ τε καὶ περιφερές ἐν αὐτοῖς; 436e1). We would then be at liberty to point out that these are different, so that the spinning tops do not present counterexamples to our principle.

The purport of this example is clear; and it is evidently for just this purport that the example of the top is introduced. The kinds of parts covered by (PGP) extend to abstract parts. Abstract parts, however, are conceptual parts. Hence, Plato designs his governing principle, unaugmented, to generate abstract parts. He therefore also embraces the conclusion that his argument generates conceptual parts. Since conceptual parts are neither organic nor aggregative parts, it also follows that this argument in its present form cannot establish the existence of such parts in the soul.

More to the point, and we now move on to our second premise (that the argument for soul-division is compatible with the form of simplicity required for the Affinity Argument of the *Phaedo* and the Special Vice Argument of *Republic* x), in establishing these forms of parts, Plato does nothing to undercut any argument for simplicity he might be inclined to offer. For the argument from simplicity assumes that necessarily, something is incomposite iff it is eternal. Even Forms, however, the very basis of the analogy in the Affinity Argument, have conceptual parts. For anything with properties has conceptual parts, and Forms have properties. So too with a point: it has conceptual parts, but not the kinds of parts required for dissolution. Such parts would include, e.g., spatial parts, which might be either organic or aggregative. It follows, then, that the sorts of parts generated by Plato's argument for soul division do not impugn any argument from simplicity he might be inclined to offer.<sup>17</sup>

I turn now to the second compatibility argument, one which abstracts slightly from the fine features of Plato's mereology. This argument takes very seriously indeed the image of Glaucon, by taking Plato at his word when he says that the soul in its true nature may not be multiform. More to the point, it establishes that Plato's argument for soul division could not establish anything about the essence of the soul. For it cannot establish anything necessarily true of the soul, and all essential properties are also necessary.

When he says that the soul in its true nature may well be uniform, Plato seems to resist the suggestion that the soul is essentially composite. Some contemporary interpreters have picked up on this, arguing that any understanding of *Republic* iv according to which it establishes an *essentially* composite soul would render it irrevocably incompatible with both the *Phaedo* and *Republic* x.<sup>18</sup> So, the dominant question has become whether the argument establishes essential composition in the soul.

This question is, however, in one way insufficiently fine-grained. For, given the distinctions we have introduced, the question put thus bluntly is at best misleading. As we have seen, an argument from simplicity is compatible with the presence of conceptual parts, although it is not compatible with the existence of aggregative parts. Indeed, the argument from simplicity is compatible with the presence of essential conceptual parts; this is, of course, welcome, since *everything* with an essence has some essential conceptual parts. That being the case, the only pressing question concerns whether the argument for soul division establishes the existence of essential aggregative parts; for these are the sorts of parts into which an entity can resolve when going out of existence. Since (PGP) does not generate aggregative parts at all, it *a fortiori* does not generate essential aggregative parts. Hence, Plato's argument for soul division does not generate essential parts of a sort incompatible with the soul's being simple in the way requisite for the Affinity and Special Vice Arguments.

That acknowledged, it is worth observing something quite important about the argument, a simple point whose neglect has generated significant and long-term misconceptions about the force of the argument and its place in Plato's development. This is that Plato's argument for soul-division *could not* establish essential parts of any kind. The reason is simple. Premise (6) (that in fact our souls sometimes have an appetite to drink yet refuse to drink) is an *a posteriori* premise. It is also contingent. (6) appeals to a naked fact of experience, to the effect that we sometimes experience psychic conflict. The conclusion, then, is that *given that conflict*, it follows that our souls have the features or elements distinguished by the argument. This, though, is a straightforwardly hypothetical necessity. Hence, none of the features established by the argument can be *de re* necessities. Since every essential feature is also a *de re* necessary feature, it follows that the argument for soul division could not establish anything essential about the soul at all. From this it follows trivially that it could not establish that the soul is essentially composite.

This argument can be expanded and strengthened by considering a tempting objection.<sup>19</sup> The point assumes that since it contains an *a posteriori* premise, Plato's argument cannot establish anything necessary about the soul. *A fortiori*, it cannot establish anything essential. That is, the point relies upon a tacit inference to the effect that proposition *p* is necessary only if it is not knowable *a posteriori*. Perhaps Plato may have assumed that such an inference would be stable, or perhaps not; he does not address the issues in just these terms. Still, we cannot accept any such assumption, because we know of the existence of *a posteriori* necessities. As Kripke has pointed out,<sup>20</sup> the *a posteriori/a priori* distinction is an epistemological distinction, while the necessary/contingent distinction is metaphysical in character. The distinctions are not the same, and indeed need not be coextensive. In fact, he contends, they are not coextensive: that water = H<sub>2</sub>O is necessary but knowable only *a posteriori*. So, and this is the force of the current objection, Plato may well and rightly have relied upon an *a posteriori* premise in an argument which draws a conclusion about the necessary features of the soul. If so, the fact that one premise in the argument is *a posteriori* does nothing to impugn the possibility of its drawing a conclusion concerning the essential features of the soul.

This objection is unconvincing. In the first instance, though widely influential, Kripke's arguments for the existence of *a posteriori* necessities were refuted already in 1983.<sup>21</sup> That said, the relevance of Kripke's arguments to the point at hand is in any case already suspect. Suppose, that is, as many do, that at least one of Kripke's arguments for the existence of the necessary *a posteriori* is sound. Even this concession would do nothing to undermine the observation that Plato's appeal to an *a posteriori* premise is incompatible with his inferring anything *de re* necessary about the soul. For Kripke's dominant arguments pertain not to *de re* necessities, but to *de dicto* necessities, and then only to a highly specific kind, those expressed by identity statements flanked by rigid designators. Plato's appeal to the fact that we indeed at times experience internal psychic conflict is not a claim about identity statements at all, and thus falls outside the scope of Kripke-style arguments for the existence of the necessary *a posteriori*. Moreover, the apparatus of rigid designation is hardly relevant to Plato's claim (6), which, again, holds: In fact, our souls sometimes both have an appetite to drink and refuse to drink (439c1-2). This phenomenological appeal to our psychological lives has nothing whatever to do with highly theoretical and specific claims concerning the behavior of rigid designators, if there are such.<sup>22</sup> Hence, the

putative existence of *a posteriori* necessities of the sort envisaged by Kripke and his followers has no bearing on the question of whether Plato's argument for soul tripartition could eventuate in essential features of the soul.<sup>23</sup>

It is possible, I suppose, that someone might want to rewrite Plato's actual argument in such a way that it would dispense with premise (6), so that the argument makes no *a posteriori* appeal at all. Such an argument might prove interesting; but it would not be Plato's argument. More to the point, it would be odd and tendentious to use any such argument to create a problem for Plato along the lines currently under consideration. To insist that Plato had contradicted himself by promoting the simplicity of the soul in one context while in another advancing an argument which might be rewritten so as to require a commitment to an essentially composite soul would really be to insist only that Plato might have contradicted himself, but did not.

Given the features of Plato's actual argument, it would be a mistake, as Plato says it would be a mistake, to infer that the forms of complexity ascribed to the soul throughout the *Republic* are, in any relevant sense, incompatible with the simple souls he understands to be our true natures.<sup>24</sup> These Platonic souls are simple souls.

#### IV. Conclusions

Although it would be incorrect to derive any sweeping unitarian conclusions from these investigations, two results remain noteworthy. First, there is no reason to doubt Plato's contention that the argument for soul-division in *Republic* iv is compatible with the conception of unitary souls upon which he relies elsewhere, mostly notably in the *Phaedo* and in *Republic* x. Second, it is incorrect to posit a shift in metaphysical psychology as an explanatory factor in Plato's (alleged) anti-Socratic affirmation of *akrasia*. For there is no reason to suppose that he underwent any such shift.<sup>25</sup>

## Bibliography

Annas, J. 1981. *An Introduction to Plato's Republic* (Oxford: Clarendon).

Archer-Hind, R. 1881. "On Some Difficulties in the Platonic Psychology." *The Journal of Philology* 10: 120-31.

Crombie, I. 1962. *An Examination of Plato's Doctrines*. Vols. i and ii (London: Routledge).

Cross, R. and Woozley, A. 1964. *Plato's Republic: A Philosophical Commentary* (London: Macmillan).

Frede, M. 1992. "Introduction." In *Plato: Protagoras*. Trans. S. Lombardo and K. Bell (Indianapolis, Ind.: Hackett).

Gallop, D. 1975. *Plato, Phaedo*. Trans. with comm. and notes (Oxford: Clarendon).

Grote, G. 1867. *Plato and the Other Companions of Sokrates*. Vols. i-iii (London: J. Murray).

Guthrie, W. 1975. *A History of Greek Philosophy*. Vol. iv (Cambridge: Cambridge University Press).

Irwin, T. 1995. *Plato's Ethics* (Oxford: Oxford University Press).

\_\_\_\_\_, ed. 1999. *Classical Philosophy* (Oxford: Oxford University Press).

Kripke, S. 1972. "Naming and Necessity." In *Semantics of Natural Language*. Ed. G. Harman and D. Davidson (Dordrecht: Reidel), 253-355.

Miller, F. 1999. "Plato on the Parts of the Soul." In *Plato and Platonism*. Ed. J. Van Ophuijsen (Washington, D.C.: The Catholic University Press of America), 84-101.

Murphy, R. 1951. *The Interpretation of Plato's Republic* (Oxford: Clarendon).

Nettleship, R. L. 1898. *Lectures on the Republic of Plato* (London: Macmillan).

Tichy, P. 1983. "Kripke on Necessary A Posteriori." *Philosophical Studies* 43: 225-41.

Vander Waerdt, P. 1985. "Peripatetic Soul-Division, Posidonius, and Middle Platonic Moral Psychology." *Greek Roman and Byzantine Studies* 26: 373-94.

Woods, M. 1987. "Plato's Division of the Soul." *Proceedings of the British Academy* 73: 23-48.

## Notes

1. Frede offers this sort of judgment in commenting on the moral psychology of the *Protagoras*: "If we find this highly intellectualist account of the passions as judgments of some kind implausible, we should keep in mind that it is only Plato, in the *Republic* (iv, 437b ff.), who, precisely to explain how one can act against the judgment of one's reason, for the first time introduces different

parts of the soul, each with its own desires, allowing us to understand how irrational desire may overcome the dictates of desire and reason. Here in the *Protagoras*, Socrates seems to argue as if the soul just were reason, and the passions were reasoned beliefs or judgments of some kind, and as if, therefore, we were entirely guided or motivated by beliefs of one kind or another. On this picture of the soul, it is easy to see why Socrates thinks that nobody acts against his knowledge or even his beliefs: nothing apart from beliefs could motivate such an action" (1999, xxx). Irwin, commenting on *Republic* iv, reaches a similar conclusion: "The division of the soul into parts with potentially conflicting desires seems to explain incontinence. The explanation rejects Socrates' dissolution of apparent incontinence into mistaken belief about comparative goods and evils. If we have desires that are not responsive to reasoning about the good, it is not surprising that they sometimes move us to action in conflict with our belief about the good" (1999, 287). Similarly, Miller (1999, 100) concludes: "Plato's early theory of the soul was substantially revised in order to explain conflicting motivations. Central to this revision were the conceptions of the soul as a self-moving principle and of desires as motions of the soul. The *Phaedo*'s doctrine of the simplicity of the soul on both counts had to yield to the tripartite psychology."

2. This point seems appreciated by Woods (1987), 45 n. 45.

3. In making this claim, I do not maintain that Plato's views about the soul undergo no form of development whatsoever. On the contrary, there seem to me some important differences between the conception of soul in the *Phaedo* and *Republic*, centering especially on its corporeal concomitants. I maintain, rather, a metaphysical thesis about Plato's mereology, that the soul is not essentially partite.

4. I do not here distinguish between eternality and sempiternity, as one must in some other contexts in Plato. Here I treat eternity as a generic notion of ceaseless existence.

5. For example, Plotinus, *Enneades* i. 1. 12, Berkeley, *Principles*, 141: "We have shown that the soul is indivisible, incorporeal, unextended; and it is consequently incorruptible."

6. So, for example, Gallop, commenting on Plato's introduction in the *Phaedo* of an opposition between an exhaustively rational soul and the emotions and desires housed in the body: "This opposition is not between 'reason' and some other 'part' of the soul, but simply between soul and body. No distinction is drawn in the *Phaedo* between 'reason,' 'spirit,' and 'appetite,' which Plato treats elsewhere as separate 'parts' of the soul—see *Republic* 435a-441c and *Phaedrus* 246a-b, 253c-e" (1975, 89).

7. So Annas (1981, 346), commenting on Plato's argument for immortality (608c-611a): "Plato does not see, or is unmoved by, the question-begging nature of this argument. He goes on (611a-612a) to a problem that this argument raises. The soul has throughout the *Republic* been treated as having parts, as being the area of psychological complexity and division. But the soul just proved immor-

tal is *unitary* . . . Plato here betrays a tendency (cf. *Phaedo* 78b-c) to think of destruction as being a breaking-up and dissolution so that anything indestructible must have no parts or internal complexity. However, if the soul has internal complexity, how can it be immortal? Plato's reply here is that the soul's true nature is immortal, and that our notion of it as being composite and liable to internal conflict is not a true view. . . . Why should we think that immortal simplicity is the soul's true nature, rather than its perceived complexity? Fantastic as it seems, Plato is laying more weight on this ridiculous little argument than on the whole of the rest of the *Republic* with its use of the composite soul." Similarly, Guthrie (1975, 555): "This is certainly no improvement on the arguments for immortality in the *Phaedo* and *Phaedrus*, and for Plato's sake one would hope that it was not very seriously meant. . . . Unless I have misunderstood it, the illogicality of the argument, apart from its astonishing premises, is patent."

8. The Special Vice Argument holds that *x* can be brought out of existence, if at all, only by the vice peculiar to it (τὸ σύμφυτον κακὸν; 609a8) and nothing else, that injustice is the special vice of the soul, but that injustice is not the sort of vice which can destroy the soul. Plato concludes, then, that "it is clear the soul is something which always exists, but if it always exists, it is immortal;" δῆλον δτι ἀνάγκη αὐτὸ δεὶ δν εἰναι· εὶ δ' δεὶ δν, ἀθάνατον; 611a1).

9. Cf. *Phaedrus* 246a, and *Tim.* 72d, where Plato is similarly agnostic or at least cautious.

10. This problem was already formulated crisply in Grote (1875, Vol. ii, 159): "In the *Republic* and *Timaeus*, the soul is a tri-partite aggregate, a community of parts, a compound: in the *Phaedon*, Sokrates asserts it to be uncompounded, making this fact a point in his argument."

11. Archer-Hind (1881, 128) formulates the difficulty this way: "We are therefore driven to choose between the following suppositions: (1) Plato has directly contradicted himself on a point of the gravest importance; (2) the term *psuchē* is used by him in different senses; (3) the expression *thnēton eidos psuchēs* is to be explained so as to harmonise with Plato's other statements on the subject." He opts for (3), as do I. Our reasons, though compatible, diverge sharply.

12. In this I agree with Guthrie (1975, 478): "when Plato wrote book 10 he had not forgotten what he said in book 4. The one change came after the *Phaedo*, where passion and appetite were attributed to the body, not the embodied soul. Essentially the soul remains what it was there, simple and akin to the divine." I do not, however, accept his reasons for this conclusion. Guthrie relies exclusively on the view of Nettleship (1898, 154). Nettleship does nothing, however, to address the question of Plato's mereology.

13. This formulation of the argument agrees with Irwin (1995, 204).

14. N.b. that it must be our *souls* which Plato tacitly accepts as the subjects of both these attitudes. If it were, e.g., our *selves*, then we would not have even a *prima facie* argument for soul division; one might well hold, e.g., the picture

commonly associated with the *Phaedo*, according to which the rational soul is the seat of one form of desire and the body the seat of another.

15. For an excellent discussion of this issue, see Miller (1999), 92-95.

16. One noteworthy study which pays attention to the character of Plato's psychic parts is Woods (1987), who is opposed to the tendency of some others, which he locates *inter alia* in Crombie (1962, Vol. i, 356), of treating the parts of the soul as mere psychic phenomena. Woods aims "to show that Plato needs more than a threefold classification of phenomena: he needs to establish distinct sources of action" (25). Even Woods does not ask what would be required, in terms of Plato's mereology, to establish distinct sources of action. In one sense, I am inclined to agree with Woods; but in this sense, the three sources will not be proper parts from which a unified soul is fused.

17. In one sense, then, I agree with Archer-Hind (1881, 125), who concludes against Zeller that the three parts of soul are "verschiedene Thätigkeitsformen," as opposed to "verschiedene Theile."

18. Thus, Cross and Wootzley (1964, 120) maintain that an essentially tripartite soul would force an "irresolvable contradiction" between *Republic* iv and x. As will be seen in what follows, in one way I agree and in another I disagree. The question turns crucially on the conception of "parts" employed in this judgment.

19. Several people have voiced this objection. Among them, Nicholas White has put it most clearly and forcefully.

20. Kripke (1971), 267-71.

21. By Pavel Tichay (1983), whose arguments I will not reproduce, but which I endorse almost completely.

22. So, although a second form of argument offered by Kripke does pertain to *de re* necessity, including especially facts about origination, these arguments are equally irrelevant to the experiential appeal in Plato's premise (6). For these arguments too rest upon an appeal to the apparatus of rigid designation.

23. That Plato did not himself understand the argument to generate three (and only three) essential features of the soul seems confirmed by his observation toward the end of *Republic* iv that the just person is the one who has harmonized his reason, spirit and appetite, as higher, lower, and intermediate parts, "as well as any others there may be between them" καὶ εὶ ἀλλα ἄττα μεταξὺ τυγχάνει δντα; 443d7). It would be perverse and cavalier to allude to other parts of the soul, evidently on par with the three delineated by the arguments of *Republic* iv, if these parts were understood as all essential. This would generate an unacceptable psychic fission. Indeed, it would seem to land Plato in the awkward dilemma clearly articulated by Miller (1999, 100): "Plato's argument for a tripartite soul thus faces a dilemma: depending on how we understand the argument, either it establishes too little (namely, only two parts of the soul) or too much (a part of the soul corresponding to every desire)." If, however, the parts in question are conceptual parts, and in any case, not essential to the soul, Plato is at liberty to generate as many parts as he requires for any given eristic

purpose. Again, it is surely noteworthy that in this context he himself finds congenial the existence of other parts. On this point see also Murphy (1951, 29), with whose view I agree.

24. As a modest confirmation of this conclusion, it is worth recalling that Plato shies away from speaking in terms of *parts* (*merê* or *moria*) in *Republic* iv, employing the term *meros* first at 442b11 already near the end of the chapter and well after the dominant argument for soul-division has been delivered. Plato prefers three forms of locutions: (i) the article with a relative clause, (ii) adjectives, and (iii) noun phrases other than part (e.g., *genê* and *eidê*, kinds or types. It is also worth recalling that the dominant interpretation of Plato in antiquity, from Aristotle (MM 1182a26) through the Middle Platonists (see, e.g., Arius Didymus 38.14), ascribed only bipartite division to Plato, suggesting either: (i) a willful misunderstanding of Plato's mereology, as Vander Waerdt (1985) would have it, or, as I prefer, (ii) a recognition that the soul division of *Republic* iv does not establish parts incompatible with other axes of division, for example, into the *alogon* and *logon echon*. No violence is done to Plato in this division, since it is compatible with the tripartite division established in *Republic* iv.

25. I thank Richard Cameron, Gabriela Carone, Richard Geenen, Rachel Singpurwalla, and Ellen Wagner for helpful discussions of these matters. I am also grateful to Gail Justin for valuable comments on an earlier draft.

## The Analogy of City and Soul in Plato's *Republic*

Bernard Williams

In making the first construction of the city, there is an assumption that it should be able to tell us something about *δικαιοσύνη* in the individual: we look to the larger inscription to help us read the smaller one, 368D. But, as Plato indeed implies, the larger inscription will help with the smaller only if they present the same message. What is Plato's reason for expecting the same message? Basically, it is that *δίκαιος* applies to both cities and men, and that it signifies one characteristic: "So the just man will not differ at all from the just city, so far as the character of justice is concerned, but will be like it" (*καὶ δίκαιος ἄρα ἀνὴρ δικαίας πόλεως κατ' αὐτὸ τὸ τῆς δικαιοσύνης εἶδος οὐδὲν διοίσει, ἀλλ' ὅμοιος ἔσται*, 435B). That there should be some kind of analogy between cities and men in respect of their being *δίκαιος* would seem to be a presupposition of asking the question "what is *δικαιοσύνη*?" and expecting one answer to it.

Indeed at 434E Plato says that when we transfer what we have said about the city back to the man, we may find that it does not work out; but the moral will be that we should go back and try again and "perhaps by looking at the two side by side and rubbing them together, we may make justice blaze out, like fire from two sticks." Plato clearly has a fair confidence that this technique will work: his confidence is in what I shall call the *analogy of meaning*.

At 435E, however, he takes what is in fact a different tack. Proceeding there to the division of the soul, he seems at first sight to be backing up the “analogy of meaning.” “Are we not absolutely compelled to admit that there are in each one of us the same kinds and characteristics as there are in the city? For how else could they have got there? It would be ridiculous to imagine that among peoples who bear the reputation for being spirited . . . the spirited character in their states does not come from the individual citizens,” etc. This looks as though it means that we call a city, people, etc. “spirited” because most or all of its individual persons can be called “spirited”—and for certain terms, this style of account is very reasonable.

But for such terms (the three examples that Plato gives at 435E correspond, it is worth noting, to the three elements of his analogy), so far from having something that backs up the previous principle of finding a common characteristic in virtue of which both cities and men are called so-and-so, we have something that defeats it. For if we say that “F” is applied to the city just because it is applied to the men, we have already explained how the term can be applied to both cities and men, and to go on from there to look for a similar explanation of how “F” applies to men is at least pointless, since the phenomenon which set off the search for the analogy in the first place, viz., the fact that “F” applies to both cities and men, has already been explained. If, moreover, the rule for applying “F” to cities is taken as itself the common *λόγος* that we were looking for, then we have not just pointlessness but absurdity, since the common *λόγος* will have to be something like “x is F if and only if x has constituent parts which are F,” which leads to a regress. Thus the argument at 435E, so far from backing up the “analogy of meaning,” defeats it.

Plato in any case does not seem to think that every term which can be applied to both cities and men obeys the rule of 435E. Thus at 419A ff. (the beginning of Book IV), answering Adeimantus’ objection that the guardians get a thin time of it, Socrates says that a city’s being sublimely happy does not depend on all, most, the leading part, or perhaps any, of its citizens being sublimely happy, just as a statue’s being beautiful does not depend on its parts being severally beautiful. This contradicts the principle of 435E, and certainly contains a truth. Leaving the importantly, and indeed deeply, contentious case of “happy,” we can certainly agree that a large crowd of sailors is not necessarily a crowd of large sailors, while an angry crowd of sailors, on the other hand, is a crowd of angry sailors. So what Plato has here are two classes of term: one class (“angry,” “spirited,” etc.) obeys the rule of 435E, which we may call *the*

*whole-part rule*; while the other class (“large,” “well-arranged,” etc.) does not.

However, Plato does not proceed along the lines of this distinction. Rather, for an indeterminately large class of terms, possibly including *δίκαιοσύνη*, he wants to say both:

(a) A city is F if and only if its men are F;

and

(b) The explanation of a city’s being F is the same as that of a man’s being F (the same *εἰδος* of F-ness applies to both).

The combination of these, as we have already seen, could lead to a regress, but Plato avoids this by holding (a) only for the city-man relation, and not for the relation of the man to any further elements—that is to say, he does not take (a) as itself identifying the *λόγος* of F-ness. Thus “F” does not occur again in the explanation of what it is for a man to be F: at that stage, it is reduced to something else. Thus the explanation of a man’s being *δίκαιος*, and the *λόγος* of *δίκαιοσύνη* in general, are alike given us by the formula

(c) Each of the elements (*λογιστικόν*, *θυμοειδές*, and *ἐπιθυμητικόν*) does its job,

which of course implies

(d) *λογιστικόν* rules.

Applying (a) to the particular case of *δίκαιοσύνη*, we get

(e) A city is *δίκαιος* if and only if its men are;

while at the same time, for a city as for a man, we have the requirement that its being *δίκαιος* consists in (c)’s being true. But what does (c) mean of a city? For like cities, the elements of cities consist of men: and how are the characters of these elements to be explained? Here it seems the whole-part rule must certainly apply—it was, we remember, with reference to these characteristics that Plato introduced us to it. We shall have

(f) An element of the city is logistic, thymoeidic, or epithymetic if and only if its men are.

But the *δικαιοσύνη* of a city, as of anything else, consists in (c)'s being true. So in order to be *δίκαιος*, a city must have a logistic, a thymoeidic, and an epithymetic element in it. Since it must have an epithymetic element, it must, by (f), have epithymetic men: in fact, it is clear from Plato's account that it must have a majority of such men, since the lowest class is the largest. So a *δίκαιος* city must have a majority of epithymetic men. But an epithymetic man—surely—is not a *δίκαιος* man; if he is not, then the city must have a majority of men who are not *δίκαιοι*, which contradicts (e).

This contradiction is, I believe, powerfully at work under the surface of the *Republic*. Remaining still at a very formal and schematic level, we get another view of it by asking what follows if we accept (e) and also take the analogy between city and soul as seriously as Plato at some points wants us to. Since the men are *δίκαιοι*, of each man (d) will be true, and *λογιστικόν* (no doubt in some rather restricted way) will be at work in each member even of the lowest and epithymetic class. Some minimal exercise of *λογιστικόν* would seem to be involved in bringing it about that each man sticks to his own business, which is the most important manifestation of social *δικαιοσύνη*: though it is very notable that Plato repeatedly uses formulations abstract and impersonal enough to prevent such questions pressing to the front. (A very striking example of this is at the point where *δικαιοσύνη* is first, after the hunt through the other cardinal virtues, pinned down. At 433C-D we have a reference to the beneficent effects of the *φρόνησις* of the guardians, but by contrast with this, “that which is in” [τοῦτο . . . ἐνόν] even slaves, artisans, women, etc., and which makes the city good, is represented not as a characteristic of theirs, but merely as a *fact*, that each minds his business [ὅτι τὸ αὐτοῦ ἔκαστος εἰς ὃν ἐπράττε καὶ οὐκ ἐπολυπραγμόνει, 433D4-5]. Clearly, this fact cannot be “in” these people—the question is, what has to be in these people to bring about this fact.)

But now if the epithymetic class has in this way to exercise some *λογιστικόν*, and this helps it to stick to its tasks, recognize the rulers and so forth, and if we read this result back through the analogy to the individual soul, we shall reach the absurd result that the *ἐπιθυμητικόν* in a just soul harkens to *λογιστικόν* in that soul through itself having an extra little *λογιστικόν* of its own. Recoiling from this absurdity, we rec-

ognize that in the individual soul, the *ἐπιθυμητικόν* cannot really harken; rather, through training, the desires are weakened and kept in their place by *λογιστικόν*, if not through the agency, at least with the cooperation, of *θυμοειδές*. If with this fact in our hand we come back once more across the bridge of the analogy to the city, we shall find not a *δίκαιος* and logically cooperative working class, but rather a totally logistic ruling class holding down with the help of a totally thymoeidic military class, a weakened and repressed epithymetic class; a less attractive picture. The use of the analogy, it begins to seem, is to help Plato to have it both ways.

Does Plato intend us to accept the proposition (e), that the citizens of the *δίκαιος* city are themselves *δίκαιοι*? The question is not altogether easy. The passage 433-4, from which I have already quoted the most notable evasion, manages to create the impression that the answer must be “yes” without, so far as I can see, ever actually saying so. An important contributory difficulty here is the point which has been often remarked, that the earlier account of *σωφροσύνη* has left *δικαιοσύνη* with not enough work to do, so that it looks like merely another way of describing the same facts. In the case of *σωφροσύνη*, he comes out and says that it is a virtue of all citizens (431E-432A); but the route to this conclusion has several formulations which make even this seem shaky (431B-D, particularly: “the desires in the many and vulgar are mastered by the desires and the wisdom in the few and superior”). The tension is always the same. The use of the analogy is supposed in the upshot to justify the supreme rule of a logistic element in the city, where this element is identified as a class of persons; and it justifies it by reference to the evident superiority of a soul in which the logistic element controls the wayward and chaotic desires. But this will work only if the persons being ruled bear a sufficient resemblance to wayward and chaotic desires—for instance, by being persons themselves controlled by wayward and chaotic desires. And if they are enough like that, the outcome of Plato's arrangements will be less appealing than first appears.

Suppose, then, we give up the proposition that all or most people in the *δίκαιος* city are *δίκαιοι*; thus we give up the whole-part rule for *δικαιοσύνη*. We might, at the same time, put in its place something rather weaker than the whole-part rule, which we might call the *predominant section rule*:

(g) A city is F if and only if the leading, most influential, or predominant citizens are F.

The effect of using (g) with δικαιοσύνη is of course to cancel any implication that the citizenry at large are δίκαιοι—it merely gives us something that we knew already, that the guardians are δίκαιοι. But the importance of (g) is in no way confined to the case of δικαιοσύνη—it is a rule which Plato appeals to often, and particularly in his discussions of the degenerate forms of city in Book VIII. It is in the light of the *predominant section* idea that we should read the reiteration of the whole-part idea which introduces those discussions at 544D. If we look at some of the things that Plato says about the degenerate cities, this will lead us back again to the just city, and to the ineliminable tension in Plato's use of his analogy.

With the degenerate cities, it is clear in general that not all the citizens are of the same character as the city, and there are references to citizens of a different character. The tyrannical city is, not surprisingly, that in which there is most emphasis on the existence of citizens different in character from the tyrant: 577C “the whole, so to speak, and the best element is dishonorably and wretchedly enslaved”; cf. 567A, 568A “the best people hate and flee the tyrant.” In other kinds of city, there may be a minority of citizens of a character inferior to that of the city as a whole: there may be a few men of tyrannical character in cities where the majority is law-abiding (575A); if few, they have little influence, but if there are many, and many others who follow their lead, then they produce a tyrant (575C). We can notice here that even in a tyranny there is a requirement that a substantial and influential section of the citizens should share the character of the city. Again, at 564D we are told that the “drones” are already present in an oligarchy, but in a democracy they become the leading element (*τὸ προεστὸς αὐτῆς*).

The democracy, however, presents a special difficulty. Plato says that the distinguishing mark of a democracy is that it is the state in which one finds men of every sort (*πάντοδαπόλ*, 557C), and like a garment of many colors it is decorated with every sort of character (*τὰσιν ἡθεσιν*, 557C.). Having said this, it would be impossible for Plato to say that all the citizens were “democratic” men as described at 561D *al.*—always shifting, without expertise in anything, prepared to indulge any ἐπιθυμία, etc. Nor should it be easy for him to say that the majority are such men. Yet this is precisely what he has to say. The “predominant section” rule says that the character of the state is derived from that of the ruling citizens. In the cases where the rulers are few, this will not necessarily imply much about the character of other citizens, for the few may hold their power by force, threats, etc. (as in the case of the tyrannical state,

already considered: and cf. 551B, the origin of the oligarchy). Plato says that a democracy will also come into being by threat of force, 557A—but this is merely vis-à-vis the ruling oligarchs. A democracy is a state in which the many rule, and if it gets its character from that of its rulers, then the majority must have a “democratic” character. This, on the face of it, sorts none too well with the claim that the democratic state will particularly tend to contain all sorts of character—the “democratic” character seems in fact to be a special sort of character. Moving between the social and the individual level once more, Plato seems disposed to confound two very different things: a state in which there are various characters among the people, and a state in which most of the people have a various character, that is to say, a very shifting and unsteady character.

These people, moreover, are the same people that constitute the lowest class in the δίκαιοις city; so we are led back once more to the question we have already encountered, of how, consistently with Plato's analogy and his political aims, we are to picture their quiescent state when λογιστικόν (in the form of other persons) rules. It may be said that in the difficulties we have found about this, we have merely been pressing the analogy in the wrong place. The essential analogy here might be claimed to be this: just as there is a difference between a man who is controlled by λογιστικόν and a man who is controlled by ἐπιθυμία, so there is such a difference between states, and to try to infer the condition of the epithymetic class when it is ruled from its condition when it is not ruled is like trying to infer the condition of a man's ἐπιθυμίαι when *they* are ruled from their condition when they are not. What we are concerned with (it may be said) is the healthy condition of man or city, and relative to that the difference between a good and a bad state of affairs can be adequately—and analogously—explained for each.

Such attempts to ease out the difficulties only serve to draw attention to them. For, first, certain things *can* be said about the ἐπιθυμίαι when they are “ruled.” For instance, there is the notable difference between a man who has his ἐπιθυμίαι under control, so that he does not act on them except where appropriate, but for whom they are nevertheless very active, so that control is the outcome of struggle and inner vigilance; and a man whose λογιστικόν has achieved inner peace. That inner peace, again, might be of more than one kind: some ἐπιθυμίαι might be mildly and harmoniously active, or there may have been some more drastically ascetic achievement—*solitudinem faciunt, pacem appellant* could apply to inner peace as well. But these differences, read back into the political case, precisely revive the earlier problems. Inner peace is what Plato

must want, but that in the political case requires the allegiance of the epithymetic element, and we are back to the question of how we are to picture that being secured. Again, a difference between the barely self-controlled man and the man of inner peace is that the first has some  $\epsilon\pi\iota\theta\mu\acute{\iota}\alpha\iota$  which the latter does not have—if a man has inner peace, then some  $\epsilon\pi\iota\theta\mu\acute{\iota}\alpha\iota$  he will have eliminated or never had. But does the difference between the good city and democracy then lie partly in the emergence in the latter of extra and more violent epithymetic persons? If so, then Plato has to explain why the working class even in the good city has to be thought of as though they were already potentially such persons. If not, we are faced with the original problem once more, of what it was in those potentially violent persons that kept them in their place in the good city.

Let us suppose that it is the inner peace model that Plato has in mind, and that it is achieved through the exercise of  $\lambda\sigma\gamma\iota\sigma\tau\iota\kappa\acute{\iota}\nu$ , on a modest scale, by the individuals in the working class. (They might have been said to possess some measure of  $\delta\varrho\theta\gamma\ \delta\acute{\o}\xi\alpha$ , if that were not inconsistent with the eccentric theory of knowledge which the *Republic* presents). If their individual  $\lambda\sigma\gamma\iota\sigma\tau\iota\kappa\acute{\iota}\nu$  helps in keeping the workers in their place, then (as we saw earlier) the analogy is no longer in full working order, since that feature cannot be read back into the soul without absurdity. But let us waive that point, and ask what has to be presupposed to keep even the remnants of the analogy going for Plato's purposes. It is not enough that in its economic function, the role of the lowest class should bear some analogy to the role of the  $\epsilon\pi\iota\theta\mu\acute{\iota}\alpha\iota$  in individual life. For if we stick merely to the nature of certain roles or functions, no argument will have been produced against the view of Plato's democratic enemy, that those roles or functions can be combined with the business of ruling. Criticism of Plato often concentrates on his opinion that ruling is a matter of expertise; but he needs more than that opinion to reach his results in the *Republic*, and has to combine with it a set of views about what characteristics and talents generally coexist at the level of individual psychology. In that area, he has to believe not only that  $\lambda\sigma\gamma\iota\sigma\tau\iota\kappa\acute{\iota}\nu$  comes in two sizes (as we might say, regular size and king size), but also that the talents and temperament that make good soldiers go with thymoeidic motivations, and the talents and temperament that make good workers go with epithymetic motivations.

Of these, the former looks plausible enough—indeed, soldierly temperament and thymoeidic motivation are perhaps necessarily connected (that is a question we shall come back to). Again, logistic supremacy and

fitness to be a guardian are of course for Plato necessarily connected. But how about epithymetic motivation and fitness to be a  $\delta\eta\mu\iota\sigma\mu\gamma\acute{\o}\varsigma$ ? Not even Plato at his loftiest can have believed that what actually qualified somebody to be a cobbler was the strength of his  $\epsilon\pi\iota\theta\mu\acute{\iota}\alpha\iota$ . The most he can have thought is that the sort of man who made a good cobbler was one who had powerful  $\epsilon\pi\iota\theta\mu\acute{\iota}\alpha\iota$ ; and this is also the least he can think, if he is to keep any of the analogy going and justify the subordinate position of cobblers by reference to their epithymetic disposition. So what we have to believe, it seems, is that cobblers are characteristically men of powerful passions—of more powerful passions, indeed, than soldiers—who nevertheless have enough rational power to recognize the superiority of philosopher kings when there are philosopher kings, but become unmanageably volatile when there are no philosopher kings.

There have been those who thought that the working classes were naturally of powerful and disorderly desires, and had to be kept in their place. There have been those who thought that they were good-hearted and loyal fellows of no great gifts who could recognize their natural superiors and, unless stirred up, keep themselves in their place. There can have been few who have thought both; Plato in the *Republic* comes close to being such a one, even though we can recognize that his heart, and his fears, lie with the first story. His analogy helps him to combine both stories, in particular by encouraging us to believe in an outcome appropriate to the second story from arrangements motivated by the first.

What about  $\vartheta\mu\o\iota\delta\acute{\o}\varsigma$  and the military class? Here there is a slightly different kind of tension in the structure. At no point, we must remember, does the structure present a simple contrast of the psychological and the political, for on both sides of that divide we have two sorts of thing: elements, and a whole which is affected by those elements. On the political side we have classes, and a state which is affected by which class is predominant among them (hence the “predominant section rule” we have already looked at); the theory is supposed to yield both an analysis and a typology of states. On the psychological side, we have “parts of the soul,” and persons in which one “part” or another is dominant; this yields, first, a classification of motives within the individual, and, second, a typology of character. The difficulties we have just been considering, about the epithymetic class, are generated across the political-psychological boundary, in the relations that Plato finds between, on the one hand, the working class and a state dominated by that class, and on the other hand, epithymetic motivation and a character dominated by such motivation. In the case of the  $\vartheta\mu\o\iota\delta\acute{\o}\varsigma$  the most interesting diffi-

culty (it seems to me) breaks out earlier, in the relations between the type of motivation that is represented by this “part of the soul,” and the type of character that is produced by its predominance. Once the type of character is established, the political consequences follow, granted Plato’s general outlook, fairly easily. Indeed, it is just the appropriateness of those consequences that seems to dictate the connection of ideas on the psychological side; whatever may be the case elsewhere in the *Republic*, here the political end of the analogy is dictating certain features of the psychological end.

I shall not attempt here any general discussion of the divisions of the soul, which, particularly with regard to the distinction between *λογιστικόν* and *ἐπιθυμητικόν*, is a large subject of great independent interest;<sup>1</sup> I shall make only some remarks about *θυμοειδές*. When it first appears, it already has a rather ambivalent role. On the one hand, it seems to be something like *anger*, and we are told, in distinguishing it from *λογιστικόν*, that it is manifested by children (441A) and animals (441B), and we are reminded of the Homeric figure who reproached his own anger. However, right from the beginning it takes on the color of something more morally ambitious (as we might put it) than mere anger or rage; the case of Leontius and other examples (439E-440E) take it rather in the direction of noble indignation, and we are told (440E) that rather than class it with *ἐπιθυμητικόν*, we should rather say that “in the strife [στάσει, a significantly political word] of the soul it takes arms on the side of *λογιστικόν*.” If *θυμοειδές* merely represented anger this would indeed be a surprising psychological claim.

The claim is indeed weakened a little at 441A, when it is said that *θυμοειδές* acts as *ἐπίκουρος* to *λογιστικόν* “if it has not been corrupted [διαφθαρῆ]: it can scarcely mean ‘destroyed’] by bad upbringing.” But the concession is not adequate. For so long as there is any conflict at all—and if there is not, the question does not arise—it clearly is possible for anger to break out, not merely against *λογιστικόν*, but on the side of *ἐπιθυμητικόν* against *λογιστικόν*. What is more interesting than that psychological platitude is the fact that Plato reveals elsewhere that he is perfectly well aware of it, and indeed in a passage where he is defending exactly the same doctrine as in the *Republic*. In the image of the chariot and the two horses in the *Phaedrus*, when the black (epithymetic) horse bolts, the white (thymoeidic) horse helps the (logistic) charioteer to bring it to a halt; and when the black horse is finally stopped, it turns on its companion and “abuses it in anger” (*μόγις ἔξ-*

*αναπνεύσας ἐλοιδόρησεν ὄργῆ*, 254C). Mere anger, Plato’s dramatic realism reveals, can always side with the devil. The thymoeidic element in the soul is from its inception more than mere anger, or indeed any other such motive which there might be good reason on purely psychological grounds to distinguish from *ἐπιθυμία* (a drive to self-destructive risk-taking, for instance).

It is to be understood, rather, by working backwards from the character which is determined by its dominance, a character which is in turn to be understood in terms of a form of life: the military or competitive form of life which it was a standard thought to contrast with the life of contemplation on the one hand and the life of gain on the other (cf. Aristotle *EN* I.5.1095b17, with, in particular, *Rep.* 581C), a contrast embodied in the Pythagorean saying about the three sorts of people that come to the Games (Iamblichus, *De Vit. Pythag.* 58). In this contrast of types of character there is also a political or social thought, of course, and that is why, as I suggested earlier, Plato has great ease in adjusting psychology and politics in the case of *θυμοειδές*: as the passage in question makes explicit (440E-441A, 440D), politics is there at its introduction. *Ἐπιθυμητικόν* has an independent psychological foundation, and Plato makes a lot of it and of its psychological relations to *λογιστικόν* in the individual, as a type of motivation. With that, I have argued, there are grave obstacles to Plato’s reading back into the city what he needs for his political conclusions, obstacles to some extent concealed by his use of the tripartite analogy.

### Note

1. For a very brief suggestion on this matter see my “Ethical Consistency,” in *Problems of the Self* (Cambridge: Cambridge University Press, 1973), 169.

## Inside and Outside the *Republic*

Jonathan Lear

### 1. Introduction

An engaged reader of the *Republic* must at some point wonder how—or if—it all fits together. There seems to be jumbled within that text a challenge to conventional justice, a political theory, a psychology, a metaphysics, a theory of education and a critique of art, music, and poetry. A brilliant work; but is it an integrated whole? A just republic, for Plato, turns out to be a harmonious, though differentiated, unity; and so the question can be rephrased: is the *Republic* a just *Republic*? Most of the illuminating discussions of the *Republic* can be seen as attempts to answer this question. I would like to suggest that this problem of unity arises in a particularly acute form for modern readers, because we are disposed to see the *Republic* as existing in bits. For we tend to conceive of psychology as the psychology of the individual. Since Plato, in the *Republic*, is concerned with the constitution of the individual psyche, it is easy for us to assume that his psychology is revealed in that account.<sup>1</sup> But this omits what, I believe, is the most distinctive aspect of Plato's psychology: a dynamic account of the psychological transactions between inside and outside a person's psyche, between a person's inner life and his cultural environment, between *intrapsychic* and *interpsychic* relations.<sup>2</sup> If we ignore these dynamic transactions, we cannot understand even individual psychology. We miss what, for Plato, holds a person together—and also what holds Plato together. For if one assumes that psy-

chology is individual psychology, the *Republic* will then look like it is composed of various bits—among them, a psychology and a political theory—and there will inevitably be a question of how they fit together. In Plato's psychology, as I understand it, this question should not arise. For *psyche*-analysis and *polis*-analysis are, for Plato, two aspects of a single discipline, psychology, which has at its core the relation between inside and outside. What holds the *Republic* together is Plato's understanding of what holds people and polis together.

In this paper I shall concentrate on two topics that lie at the heart of the *Republic*. First, there is the analogy between city and psyche. Plato thought that there were important structural isomorphisms between city or polis and psyche, and thus that he could use discoveries about one to prove results about the other. It is now widely accepted that Plato uses this analogy to fudge his arguments. Plato, so the charge goes, uses a vague analogy fallaciously, and he is thereby able to hide a fundamental tension which underlies his ideal polis. That is, he disguises the repressive relation between the ruling class and the ruled by an illegitimate comparison with the structure of the psyche. I shall argue that these criticisms look valid because Plato's psychology is not well understood.

Second, Plato's critique of the poets has inspired a wealth of deep and imaginative discussion,<sup>3</sup> but all of it has tended to concentrate on two questions: what is the effect of poetry on us?; and what is the moral value of art? Plato's argument is intriguing because, roughly speaking, we tend to think that art is good for us, while Plato argues that it is bad. Modern psychoanalysts and psychologists often think that art offers a kind of psychic salvation; while Plato treats acquaintance with Homer and the great Greek tragedians as a psychological catastrophe. And so we are led, like bees to nectar, to find a flaw in Plato's argument or, less often, to reevaluate our own aesthetics. Perhaps it is this fascination which has blinded us to the fact that we have been living on a restricted diet of questions. There are other questions, central to Plato's psychology, which as far as I know have not been asked, let alone answered. For example: who, psychologically speaking, are the poets? What, from a psychological point of view, are the poets doing in making poetry? And what is Plato doing, psychologically speaking, in banishing the poets? These are questions which, I think, tend to be obscured by assuming psychology to be the study of the individual psyche, but they come to the fore when psychology is taken to span across the boundary of an individual's psyche. For we will then see poetry as coming from some psyches and entering others, and the question naturally arises: what, from a psy-

chological perspective, is going on?

My hope is that the discussions of the polis-psyche relation and of poetry will illuminate the approach to Plato's psychology that I am advocating, and help to confirm it. As a by-product, I hope we shall also see the *Republic* as more unified than it is often taken to be.

## 2. Internalization

At the beginning of Book II, Socrates takes up the challenge, which will occupy the rest of the *Republic*, to describe justice and injustice as an "inherent condition inside the psyche" (*τίνα ἔχει δύναμιν αὐτὸ καθ' αὐτὸ ἐνὸν ἐν τῇ ψυχῇ*).<sup>4</sup> Although he proposes to look first at justice writ large in the polis,<sup>5</sup> in fact Socrates turns almost immediately to the psyche. For he begins his construction of the ideal polis with a discussion of the education of young children. And he justifies this by saying that "the beginning of any project is most important, especially for anything young and tender. For it is then that *it takes shape and any mold one may want can be impressed upon it*" (*πλάττεται καὶ ἐνδύεται τύπος, ὃν ἂν τις βούληται ἐνσημήνασθαι ἐκάστω*).<sup>6</sup> If we carelessly allow children to hear any old stories, he says, they may "take into their psyches" (*λαμβάνειν ἐν ταῖς ψυχαῖς*) beliefs that are contrary to those they should hold as adults.<sup>7</sup> Nursemaids and mothers must be allowed only to tell certain stories to their children and so "shape their psyches" (*πλάττειν τὰς ψυχὰς*). Children should not be allowed to hear the classic tales of warring gods because the young cannot distinguish what is allegorical from what is not, and the opinions they form at that age tend to be unalterable.<sup>8</sup> For Plato, humans enter the world with a capacity to absorb cultural influences. The young human psyche is like a resin, able to receive the impress of cultural influences before it sets into a definite shape. And it is clear that, for Plato, the stakes are high. The goal of achieving a well-governed polis depends on there being no one in the polis either asserting or hearing any tales which suggest that God is the cause of anything bad.<sup>9</sup> Plato believes these tales will shape the character of the future citizens.<sup>10</sup> Mothers must not be allowed to terrify their children with bad tales about gods sneaking about in disguise, "for at the same time as they blaspheme the gods, they make their children cowardly."<sup>11</sup>

If, for example, one is an honor-loving person, one should be brought up on stories of brave men doing brave deeds so as to fear slav-

ery more than death;<sup>12</sup> one should be allowed to play at and later imitate only the deeds appropriate to a guardian “lest from imitation they take (ἀπολαύω) the reality”;<sup>13</sup> one should be brought up in a rigorous program of music and gymnastics that reinforce the honor-loving part of one’s psyche;<sup>14</sup> and taken out even as a youth to observe battles;<sup>15</sup> so that when one is grown, it is through the activities of guardianship that one achieves happiness.<sup>16</sup> If this program of education and culture is successful, the qualities appropriate to guardianship should “settle into one’s character and into one’s nature” (εἰς ἔθη τε καὶ φύσιν καθίστανται).<sup>17</sup> Plato seems to be saying that through proper imitations from youth, one actually constitutes oneself as a certain type of person. Whether one develops into a noble and brave person, at one extreme, or a base coward, at the other, depends significantly on the myths one has heard from youth, the education one has received, the models one has been given to imitate. Leaving divine inspiration aside, Plato thinks that were it not for this training, one would not develop the character or nature of a guardian.<sup>18</sup>

The *Republic* is a study in the health and pathologies of cities and psyches. And the conditions of city and psyche are interdependent. The variety of pathologies of the psyche, for example, depends on the person taking in pathological structures from the culture. Culture penetrates so deeply, that a fractured polis will produce a fractured psyche. For Plato, it is only the ideal polis that can properly be called a polis or a city.<sup>19</sup> Other actual cities or poleis are only apparently such. In fact, each lacks sufficient internal unity to count as a polis: each is, in truth, many poleis or, more properly, polisparts.<sup>20</sup> But, Plato argues, for every pathological polis there is a corresponding pathology of the psyche.<sup>21</sup> The conclusion of the syllogism is that a pathological psyche is not, in fact, a psyche, but various psychic parts. So, for example, just as an oligarchy is not a polis, but two parts, a rich part and a poor part,<sup>22</sup> so an ‘oligarchical psyche’ is in fact two psychic parts: a ruling part and a ruled.<sup>23</sup> For Plato, there is not sufficient integration in the functioning of the parts for them to count as a genuine unity, a psyche. Indeed, even among the oligarchical person’s appetites there will be division and faction.<sup>24</sup> Being thrifty and acquisitive, the oligarchical person will satisfy only his necessary appetites and “enslave” his other appetites.<sup>25</sup> Because of his “lack of culture,” his unruly and unnecessary appetites spring up in him, but they are “forcibly restrained” (βίᾳ κατέχει) by the better part.<sup>26</sup> The oligarchical person is, says Plato, διπλοῦς τις, someone double.<sup>27</sup> For Plato, being double is a way of not being an integrated person: it is a divided and conflicted

existence.<sup>28</sup> In fact, the pathologies of psyche Plato examines turn out, strictly speaking, to be studies in the failures to become a psyche.<sup>29</sup>

By now it should be clear that, for Plato, satisfying the human need for culture is a process of taking cultural influences into the psyche. Let us call this process, whatever it is precisely, *internalization*. Although Plato did not have an articulated theory, he did think that imitation (*mimesis*) was a paradigmatic means of internalization. It is youthful imitations which settle the shape of one’s character and nature.<sup>30</sup> That is why musical education is preeminent: “because rhythm and harmony permeate the inner part of the psyche (καταδύεται εἰς τὸ τῆς ψυχῆς), bring graciousness to it, and make the strongest impression, making a man gracious if he has the right kind of upbringing; if not, the opposite is true.”<sup>31</sup> And it is clear that Plato thought that internalization was a largely unconscious process. Guardians should not be brought up among images of evil lest they “little by little collect all unawares a great evil in their own psyche” (κατὰ σμικρὸν . . . λανθάνωσι κακὸν μέγα ἐν τῇ αὐτῶν ψυχῇ).<sup>32</sup> One cannot change the modes of music, Plato says, without upsetting fundamental constitutional laws;<sup>33</sup> and it is clear that the causal route of this destabilization proceeds via internalization. For lawlessness, Plato says, easily creeps into music without our noticing and, “having little by little settled in there it flows into the characters and pursuits” of people (κατὰ σμικρὸν εἰσοικισαμένη ἡρέμα ὑπορρεῖ πρὸς τὰ ἔθη τε καὶ τὰ ἐπιτηδεύματα).<sup>34</sup> And so, in our education and rule of children, one should not let them be free until “a constitution is set up inside them just as in the polis” (ἐν αὐτοῖς ἀσπερ ἐν πόλει ποιεῖται καταστήσωμεν).<sup>35</sup>

For Plato, we are culture-vultures: we ‘feed’ our psyches by internalizing cultural influences. That is the psychological point of culture; and it is why education and upbringing, on the one hand, and the shaping of culture, on the other, play such a predominant role in the *Republic*. It would seem, then, that *internalization is a fundamental psychological activity*.<sup>36</sup> The fact that we are so dependent on internalization for our psychological constitution, makes us susceptible to cultural luck. Our ultimate dependency is manifest in the fact that we internalize these influences before we can understand their significance. We are dependent on culture for the constitution of our psyches, but on what does culture depend? How is culture itself shaped and formed?

### 3. Externalization

Plato suggests that culture is formed by an inverse process of psychological activity, moving outwards from psyche to polis. For example, Plato says, “there must be as many types of character among men as there are forms of government. Or do you suppose that constitutions spring from the proverbial oak or rock and not from the characters of the citizens (ἐκ τῶν ἡθῶν τῶν ἐν ταῖς πόλεσιν), which as it were, by their momentum and weight in the scales draw other things after them.”<sup>37</sup> And character, Plato says elsewhere, is inherent in the psyche.<sup>38</sup> The same forms, he says, will be found in the polis *and in the individual psyche* (τὰ αὐτὰ ταῦτα εἰδη ἐν τῇ ψυχῇ ἔχοντα),<sup>39</sup> and the shape of the polis has to be understood as deriving from the shape of the psyche:

we are surely compelled to agree that the same forms and character-types are in each of us just as in the polis (τὰ αὐτὰ ἐν ἐκάστῳ ἔνεστιν ἡμῶν εἰδη τε καὶ ἡθη ἀπερ ἐν τῇ πόλει). They could not get there from any other source. It would be ridiculous if someone supposed that spiritedness has not come to be in polis from individuals who are reputed to have this quality . . . or that the same is not true of the love of learning . . . or the love of money.<sup>40</sup>

It would seem, then, that for a significant range of psychopolitical predicates F,

(EK) If a polis is F, there must be some citizens whose psyches are F who (with others) have helped to shape the polis.

This is easiest to see in the case of the just polis.<sup>41</sup> It will be shaped by the philosopher-king, whose thoughts are directed towards realities.<sup>42</sup> And though he will try to shape the city according to a divine paradigm,<sup>43</sup> he does so by first imitating these eternal realities fashioning himself as far as possible in their likeness (ταῦτα μιμεῖσθαι τε καὶ ὁ τι μάλιστα ἀφομοιοῦσθαι).<sup>44</sup> It is by associating with the divine order that the philosopher himself becomes ordered and divine, insofar as that is possible for humans.<sup>45</sup> The philosopher, Plato suggests, has a paradigm of the internal realities inside his psyche (ἐν τῇ ψυχῇ ἔχοντες παράδειγμα).<sup>46</sup> Although there is no existing ideal polis on earth—and thus no ideal cultural template to internalize—there is a paradigm of it in heaven, and a person studying it can constitute himself (έαυτὸν κα-

τοικίζειν) its citizen.<sup>47</sup> Only after the philosopher has shaped his own psyche by internalizing divine order is he then able to shape the polis according to what has now become the order in his psyche.<sup>48</sup>

Let us call *externalization* the process, whatever it is, by which Plato thought a person fashions something in the external world according to a likeness in his psyche. Then, for Plato, the polis is formed by a process of externalization of structures within the psyches of those who shape it. And, more generally, *externalization is a basic psychological activity*. For Plato suggests that cultural products in general are externalizations. Good rhythm, harmony, and diction, for example, should follow and fit good speech (εὐλογία); and speech, in turn, follows and fits the character of the psyche (ὁ λόγος . . . τῷ τῆς ψυχῆς ἡθει ἔπειται).<sup>49</sup> In painting and all artistic works, weaving, embroidery, architecture, the making of furniture, harmony and grace are closely related to and an imitation of good character (ἀγαθοῦ ἡθούς, ἀδελφά τε καὶ μιμήματα).<sup>50</sup> And character, as we have seen, is inherent in psyche.

Notoriously, Plato believes that education must begin by telling children false tales.<sup>51</sup> These myths are distinguished from unacceptable myths and legitimated, first, because there is truth in them,<sup>52</sup> but, secondly, because that truth is a reflection of a truth in the poet's psyche. A falsehood which is merely a falsehood in words (τό γε ἐν τοῖς λόγοις [ψεῦδος]) “is an imitation of something in the psyche, a later reflection,” (μιμημά τι τοῦ ἐν τῇ ψυχῇ ἔστι παθήματος καὶ ὕστερον γεγονός εἰδωλον) which is therefore not completely untrue.<sup>53</sup> It is precisely because this merely verbal falsehood is an externalization of something true within the poet's psyche, that it can be used, with caution, as a medicine.<sup>54</sup> By contrast, falsehood in the psyche, falsehood taken as truth (ώς ἀληθῶς ψεῦδος), is what people hate most of all.<sup>55</sup> This is ignorance in the psyche (ἡ ἐν τῇ ψυχῇ ἀγνοια). Though Plato does not say so explicitly in this paragraph, the implication is that unacceptable myths and poems are externalizations of this real falsehood (τὸ τῷ ὄντι ψεῦδος).

And so it seems that in the ideal polis, after we internalize our cultural roles by a process of education, we then externalize them in our social roles. It is by a process of internalization and externalization that we are able to conform to the rule of each performing his own task. Incoherence is avoided because Plato's is a developmental psychology. Internalization is primarily going on in unformed youths; externalization is going on primarily in adults who have already formed themselves

through prior cultural internalizations. Psyche and polis are mutually constituted by a series of internalizations and externalizations, with transformations occurring on both sides of the border.<sup>56</sup> We tend to think of the economic model in psychology as concerned with the distribution of a fixed quantity of energy—and, indeed Plato lends support to this model since he believes that when a person's desires incline strongly towards something, they are correspondingly weakened for other things.<sup>57</sup> However, if we consider Plato's psychology as a whole, it would seem that a more promising economic model would be of trade across a border. Plato's psychology is basically one of interpsychic and intrapsychic trade. What is being traded across a boundary is not unformed energy, but psychological products. They are crafted both outside and inside an individual's psyche and they are traded back and forth across the boundary of the psyche. Once inside, they become citizens of a more or less federated republic and are subject to the vicissitudes of intrapsychic conflict, before being externalized again across the border.

Plato decides first to look for justice writ large in the polis because, he says, he will then be able to read the small print of the individual psyche.<sup>58</sup> By now it should be clear that he is not relying on a mere analogy of polis and psyche, but on an isomorphism which must hold due to the way we function psychologically. Psyche and polis, inner world and outer world, are jointly constituted by reciprocal internalizations and externalizations; and the analogy is a by-product of this psychological dynamic.

#### 4. The Analogy of Psyche and Polis

One way to see the virtue of an interpretation is to see how the *Republic* looks without it. In his classic essay “The Analogy of City and Soul in Plato's *Republic*,” Bernard Williams offers the most penetrating critique we have of Plato's analogy.<sup>59</sup> According to Williams, Plato's argument is incoherent, and the analogy disguises a fundamental tension in his account of psyche and polis. If Williams is right, the *Republic* is a brilliant mess. In this section, I would like to try to rescue the analogy from Williams' critique by attending to the psychological principles which underlie it.

The analogy, for Williams, is founded on two principles. First, there is the *whole-part rule*:<sup>60</sup>

(a) A city is F if and only if its men are F.

Second, there is the *analogy of meaning*:<sup>61</sup>

(b) The explanation of a city's being F is the same as that of a man's being F (the same *eidos* of F-ness applies to both).

Although it appears that these two principles support each other, Williams argues that is not so: the whole-part rule in fact “defeats” the analogy of meaning:

For if we say that “F” is applied to the city just because it is applied to the men, we have already explained how the term can be applied to both cities and men, and to go on from there to look for a similar explanation of how “F” applies to men is at least pointless, since the phenomenon which set off the search for the analogy in the first place, viz. the fact that “F” applies to both cities and men has already been explained. If, moreover, the rule applying “F” to cities is taken as itself the common *logos* that we were looking for, then we have not just pointlessness but absurdity, since the common *logos* will have to be something like “x is F if and only if x has constituent parts which are F”, which leads to a regress.<sup>62</sup>

However, Plato does not in fact think that F is applied to a polis “just because” it is applied to its citizens. Even if he were committed to principle (a) (or some variant),<sup>63</sup> the principle cannot fully capture Plato's intentions. For the principle describes a formal relation between polis and citizens, whereas Plato believes the formal relation holds in virtue of causal-psychological transactions. Plato's point (at 435E) is not that a spirited polis, say, is spirited simply in virtue of having spirited citizens, but in having spirited citizens who are successful in shaping the polis in their image. And so, one has not “already explained” how spiritedness can be applied to both polis and psyche. Plato has not yet given us the explanation: he is showing us where to look for one. He is saying that there is an externalizing psychological relation from citizen to polis. The explanation of what it is that makes either polis or man spirited lies in the future. So far Plato has only given us the methodology of a research project, one based on his psychology. If this is a general point which holds for significant psychopolitical predicates, it is not pointless to move from an explanation of, say, justice in the polis to an explanation of justice in the psyche. If a just polis is an externalization of just citizens who shape it, it would be reasonable to work one's way backwards down

this externalization to learn about the psyches of these citizens. This reasoning can occur before one has any idea what the structure of justice is.

To be sure, Socrates does say that a just person and a just polis will be alike in respect of the form of justice; and he defends this claim by appeal to a semantic principle: "things called by the same name are alike in respect to that to which the name applies."<sup>64</sup> This is the basis for Williams' principle (b). Yet even if Socrates accepts this semantic principle, there remain questions about it: e.g., why should such a semantic principle hold?; why does it hold in the realm of psychopolitical predicates?; given that it does hold, how could it be legitimate to call a certain sort of person and a certain sort of polis just? Again, the semantic principle is the beginning not the end of a research project. Only a few sentences after he introduces it, Socrates explains that a wide range of political characterizations of the polis are to be understood as externalizations of the same qualities from within the psyches of the historically significant citizens.<sup>65</sup> I read this not simply as making a psychological-causal point about the relation of the polis to its citizens, but also as providing a psychological grounding of the semantic principle, at least within the range of psychopolitical predicates. The semantic principle is introduced in the course of a dialectical inquiry, and it therefore remains open to further explication and defense. It also remains vulnerable to future emendation and revision. It should not be treated as an obvious axiom forever beyond criticism or inquiry. The psychological principles of internalization and externalization help us to understand why the semantic principle might hold in spite of the fact that there are a range of predicates which apply both to polis and to psyche.

Principles (a) and (b) do not, therefore, give us Plato's reason for thinking there to be an isomorphism between polis and psyche. The isomorphism depends on psychological relations Plato believed to hold between inside and outside. If justice, for example, can be found outside (in the polis) it must have come from inside (i.e., it must be a causal outcome of just men shaping the polis according to their conception of justice). Given the psychologically dynamic relations between inside and outside, a weak version of a whole-part rule will follow as a corollary.<sup>66</sup> And so, there is neither regress nor absurdity in Plato's argument, for there is no reason to think that he has thus far given us the common *logos*. It is often thought that Plato uses his analogy to derive his psychology: that by simply claiming the analogy and looking at the structure of the polis, he derives his psychology. But once we see that psychology is not just individual psychology, we can see that the situation is pretty

much the reverse: his psychology is used to legitimate belief in isomorphism.

Williams thinks that there is a "contradiction . . . powerfully at work under the surface of the *Republic*."<sup>67</sup> The contradiction lies in the fact that if we apply principle (a) to the case of a just polis we get that

(a') a polis is just if and only if its men are

but a just polis will have a majority of appetitive (epithymetic) persons, who, by the analysis of justice, ought to be doing their proper jobs. But an appetitive person is not a just one; and that must contradict (a'). By now it should be clear that Williams is not entitled to attribute (a) to Plato, but at most

(a'') If a polis is F, then some of its men are F<sup>68</sup>

and so he is entitled to derive not (a') but

(a''') If a polis is just, then some of its men are just;

and this generates no contradiction.<sup>69</sup>

But it is clear that Williams thinks there is a contradiction here which goes beyond the validity or invalidity of this formal argument. For, he reasons, the appetitive (epithymetic) class must exercise some reason (*logistikon*) of its own, even if it is only in the service of obeying its rulers, sticking to its tasks, etc.

But now if the epithymetic [appetitive] class has in this way to exercise some *logistikon*, and this helps it stick to its tasks, recognize the rulers and so forth, and if we read this result back through the analogy to the individual soul, we shall reach the absurd result that the *epithymetikon* [appetitive part] in a just soul harkens to the *logistikon* in that soul through itself having an extra little *logistikon* of its own. Recoiling from this absurdity, we recognize that in the individual soul, the *epithymetikon* cannot really harken; rather, through training, the desires are weakened and kept in their place by *logistikon*, if not through the agency, at least with the co-operation of *thumoeides* [the spirited part]. If with this fact in our hand we come back once more across the bridge of the analogy to the city, we shall find not a *dikaios* [just] and logically co-operative working class, but rather a totally logistic ruling class holding down with the help of a totally thymoeidic military class, a weakened and repressed epithymetic class; a less attractive picture.

The use of the analogy, it begins to seem, is to help Plato to have it both ways.<sup>70</sup>

Plato's commitment to the analogy, according to Williams, forces him into absurdities both within the realm of politics and of psychology. That is the way it will look if one takes the analogy to be *merely* an analogy. If, by contrast, we view the isomorphism as a manifestation of internalization and externalization, it seems we can use the 'analogy' to form a clearer idea of how Plato understood psychological structure. This is important because Plato identifies the distinct parts of the psyche via each part's ability to enter into fundamental conflictual relations with the other parts.

Psychological structure is delineated most obviously in intrapsychic conflict. The question then is: how are we to understand psychological structure in the absence of conflict? Instead of assuming we know what psychic parts are and using the analogy to derive absurdities, let us use Plato's principles of internalization and externalization to try to find out more about what it is to be a psychic part. In the just polis, the appetitive class does have to exercise some reason of its own, to stick to its tasks, recognize its rulers, and so forth. What intrapsychic condition (of a member of the appetitive class) might have this socially harmonious behavior as an externalization?<sup>71</sup>

Plato believes this requires a certain type of intrapsychic harmony appropriate to an appetitive person. This requires that the appetitive part of his psyche harken to reason in that psyche. The question is how one might avoid the absurdity of the appetitive part needing to have a little extra *logistikon* of its own. Not surprisingly, we need to understand the psychic part as having been formed by previous internalizations. Plato, as is well known, divides appetites into necessary and unnecessary.<sup>72</sup> The necessary appetites are either unavoidable (e.g., for basic nourishment) or they are for things which do us some good. Unnecessary appetites, by contrast, are both avoidable by proper training from youth and they lead to no good (or even to bad). In an ideal polis, then, an appetitive person will be brought up so as not to have unnecessary appetites. That is why, in contrast to his pathological cousin, the oligarchic man, he does not need to hold them down by force.<sup>73</sup> Due to his education, there is nothing in him which requires forcible restraint. Such a person will only have appetites for the bare necessities of life and for things which genuinely do him good. In the well-ordered polis, Plato says, each class will enjoy the happiness that suits its nature.<sup>74</sup> Assuming that the things that do a person good are the things that give him the happiness that suits his na-

ture, in Plato's vision<sup>75</sup> the appetitive person in a well-ordered polis should have just those appetites which the polis gives him the opportunity to satisfy.

The appetitive part has thus been shaped to be responsive to reason in the psyche. The idea that appetite needs extra reason of its own derives from the thought that appetite "cannot really harken"; and this thought in turn flows from taking the conditions in which the psychic parts are isolated to be the essential conditions in which they must operate. We identify the appetitive part by seeing it functioning in opposition to reason. If this is the way it must operate then, of course, appetite cannot harken to reason. And one can be tempted to make this inference by the thought that this is the way appetites must be.<sup>76</sup> On this picture all domination by reason would ultimately have to be repression, and Plato's alleged distinction between the oligarchical person and the appetitive person in the just polis will look like propaganda.

Moreover, if a psychic part must be the way it is when it is originally isolated, it is natural to identify appetitive persons with the appetitive parts of their psyches. For since, on this assumption, the appetitive part can have no real commerce with the other psychic parts, there seems to be no other option for appetitive persons than to be driven by their appetites. This conflicts with the claim Plato makes about the difference between the oligarchical psyche and that of the appetitive person in the just polis; but again this will look as if it is Plato's problem. However, once we recognize internalization and externalization as basic psychological activities, we can see that the psychic parts can be shaped, and thus that the conditions under which we first identify them need not be the conditions under which they operate. This allows us to see that an appetitive person need not simply be someone driven by the appetitive part. And once we see that psychic parts need not always be functioning in the conflictual ways in which they are first identified, we can then grant culture a greater role in psychic formation than would otherwise be thought possible.

Consider, for example, the appetitive person or money-lover: how did his appetites ever come to love *money*? Money is the paradigm cultural artifact: it has no existence *hors de culture*. So if the appetites can be directed onto money, it would seem that culture can permeate and inform the lower elements of the psyche.<sup>77</sup> The appetitive personality will organize his personality around his appetites; and a paradigm, for Plato, is the money lover who devotes himself to the pursuit of wealth: reason will be directed instrumentally toward figuring out ways of satisfying

this desire, he will feel honor in achieving wealth-related goals, and there is a peculiar pleasure in achieving them.<sup>78</sup> The pursuit of wealth, then, is setting the overall agenda for this person's projects, and honor and reason are disciplined to serve this outlook. But within this schema there is room for the oligarchical personality, the democrat, the tyrant and (as I shall argue) the poet, all of whom are appetitive types. 'Appetitive' is thus a genus of personality organization and the variety of species is due to the fact that internalization can inform the appetitive part of the psyche.<sup>79</sup>

It might at first seem paradoxical that, on the one hand, the appetitive part is the ruling principle of an appetitive person,<sup>80</sup> while, on the other, the appetitive person should believe along with everyone else that reason should rule.<sup>81</sup> Plato is trying to have it both ways, but, within the framework of his psychology, he can get away with it. The appetitive person thinks that the peculiar pleasures available to his way of life are the best,<sup>82</sup> and, since the appetitive part rules in his psyche, his reason will be directed towards figuring out ways to secure those pleasures. But given that this appetitive person has been brought up to have just the appetites which the well-ordered polis can satisfy, his reason ought to be telling him that the best way to satisfy his appetites is to harken to the reason manifest in the laws of the philosopher-king.

In the temperate polis, Plato says, the same belief about who should rule will be inside both the rulers and the ruled (ἡ αὐτὴ δόξα ἔνεστι τοῖς τε ἀρχούσι καὶ ἀρχομένοις).<sup>83</sup> This belief helps to constitute the reason of the appetitive person in the just polis. Ironically, it is because the reason in his psyche is subservient to the appetitive part that the appetitive person submits himself to the rule of reason in the polis. Just as the appetitive person will abjure junk food for healthy bread and relishes, so he will abjure junk bonds for municipal bonds. And all the while he will be telling himself, correctly, that this is the really good investment for himself and his family. This is how the appetitive person's role in a well-ordered polis looks from an appetitive perspective. On the one hand, his reason is focused on securing gain; on the other he concludes that the best way to do this is by following the rule of reason in the polis. This would not have been possible if he had not been brought up in such a way as to internalize appropriate cultural influences and get rid of unnecessary appetites. Yet for all that he remains basically an appetitive type: organizing his life, values, and thoughts around production and acquisition. For him, justice is basically a matter of doing and getting one's own.<sup>84</sup> Temperance in the polis is like "a certain harmony" which

"spreads throughout the whole."<sup>85</sup> But if temperance spreads throughout the whole, it must spread through the whole of the whole. That is, there would not be genuine harmony in the polis if the psyche of an appetitive citizen were at war with itself. Plato does not believe the appetitive person has the *virtue* of temperance, but in a well-ordered polis, due to well-crafted internalizations, such a person will be well disposed to temperance, both inside and outside himself.

So too for the honor-loving members of society: Each will commend the distinctive pleasures of the honor-achieving life as the best,<sup>86</sup> and will try to organize his life and character around this pursuit. In a just polis, honor-lovers will be educated to hold fast to the laws, and to fear only those things which the lawgivers think are fearful.<sup>87</sup> These people will be brought up to be soldiers: they will be educated so as to be free of unnecessary appetites and to have their other appetites disciplined to the pursuit of honor. Their reason too will be directed towards honor, but they will have been educated so as to understand that the way to achieve true honor is to defend and safeguard the law (laid down by the philosopher-rulers).<sup>88</sup> Therefore, although honor is the fundamental principle of this person's life, on that very account he will, when brought up in a just polis, believe that reason should rule. Whatever one thinks about Plato's prescription for attaining health, one must, I think, acknowledge that his conception of a healthy, harmonious psyche is not just a dodge to cover up an irresolvable tension, but a natural consequence of his psychology.

The analogy between polis and psyche is a manifestation of the fact that there are important structural similarities between interpsychic relations and intrapsychic relations. But, for Plato, these structural similarities are themselves a manifestation of important psychological transactions, back and forth, between interpsychic and intrapsychic. This is true in sickness as in health. If we examine Plato's tale of political decline, we see that the degeneration occurs through a dialectic of internalization of pathological cultural influences in individuals which provokes a degeneration in character-structure (as compared to the previous generation) which is in turn imposed on the polis, which thus acquires and provokes deeper pathology.<sup>89</sup> Plato does not merely want to show that the same neurotic structure can exist in both psyche and polis, but that the pathology in each helps to bring about pathology in the other. This has not been easy to see, I suspect, because Plato's conception of pathology is not well understood.

It is, for example, easy to read his accounts of the rise of the democratic polis and the emergence of democratic man as two parallel ac-

counts which have only a structural analogy in common. In fact, Plato traces a sophisticated interaction between polis and psyche that helps to account for both. Consider, for example, Plato's account of the rise of democratic man.<sup>90</sup> He emerges from an oligarchic family, the values and goals of that family being set by the father who is himself a manifestation of an oligarchic personality. The oligarchic father is thrifty and frugal; he has organized himself around the pursuit of wealth, and tries to instill this same structure in his family.<sup>91</sup> He has been able to keep his unnecessary appetites in check, but because he has not had a proper upbringing, because he has not experienced or internalized true culture, these appetites must be held in place by the only means available to him: brute force. This is a man whose personality is held together by forcibly holding down an inner world of unruly appetites. He presents a good face to the world, but in fact exists in two bits.<sup>92</sup> The emergence of the democratic man is, roughly speaking, the return of the repressed in the next generation.<sup>93</sup> The oligarchic father creates in his family and immediate social environment a microculture, a template for internalization, which embodies contradictory demands. On the one hand, there is the demand *inside* his family for frugality so as to accumulate wealth. There is some suggestion that this demand on its own is self-contradictory. For to pursue wealth is to organize the family around the appetites; and Plato does say there is a tendency to spoil the children.<sup>94</sup> Yet to insist on frugality is to hold those appetites in check. The appetites are thus simultaneously encouraged and forcibly restrained. The only way the father knows how to instill frugality is by force. Having failed to internalize a more harmonious psychic structure, forcible restraint is the only means at his disposal: and he imposes it on his family as well as on himself. Thus the child is brought up in a miserly fashion without real education.<sup>95</sup> But, on the other hand, the oligarchical father encourages prodigality *outside* his family.<sup>96</sup> By lending others money and encouraging wastefulness, he hopes eventually to acquire their property. These people, made poor, will eventually revolt and usher in democracy.<sup>97</sup>

Here we see how the oligarchic father, by pursuing his own ends, recreates on the interpsychic stage of his family and immediate social environment a model of his own intrapsychic relations. His son, having his appetites both encouraged and held down, becomes an interpsychic correlate of the appetites within the father. However, as a member of the outer world, the son is open to other polis influences. The oligarchical father encouraged prodigality outside the family, but Plato's point is that this prodigality cannot, finally, be kept outside. The prodigal youths, en-

couraged by the oligarch, are an externalization and interpsychic correlate to the unnecessary appetites within the oligarch's psyche. Because the son's appetites have been both encouraged and held back, he is susceptible to appetitive influences around him. "Just as the city changed when one faction received help from like-minded people outside, so the young man changes when help comes from the same type of appetites outside to one of the factions within himself."<sup>98</sup> But these appetites outside are also offspring of the father. It is these appetites—whose pedigree goes back to the father—which are reinternalized in the intrapsychic battle within the son. For a while, a struggle rages both inside and outside his psyche. The father lends his influence to aid the internalized repressing forces; the young thugs on the block egg the appetites on.<sup>99</sup> But this is a struggle which the appetites have to win, because this youth never had the opportunity to internalize good cultural structures. When the appetites come knocking on the door of his psyche, they find no one is at home.<sup>100</sup> The psyche is easily reshaped, and a "democratic man" is born.

There is a problem, though, about the relation of the democratic polis and the democratic man. The democratic polis is one which contains every sort of character, like a garment of many colors.<sup>101</sup> However, as Williams points out, the democratic man is described as always shifting, following the appetite of the moment, without any expertise.<sup>102</sup> And here Williams tightens the noose:

A democracy is a state in which the many rule, and if it gets its character from that of its rulers, then the majority must have a "democratic" character. This, on the face of it, sorts none too well with the claim that the democratic state will particularly tend to contain all sorts of character—the "democratic" character seems in fact to be a special sort of character. Moving between the social and the individual level once more, Plato seems disposed to confound two very different things: a state in which there are various characters among the people, and a state in which most of the people have a various character, that is to say, a very shifting and unsteady character.<sup>103</sup>

Surely a society of many colors does not require that each of its members be a patchwork quilt. Have we finally reached the true absurdity of Plato's analogy? I don't think so. That a polis allows and even prides itself on the fact that it has various sorts of character<sup>104</sup> is, for Plato, a manifestation of the fact that it does not have a firmly established sense of better and worse. There can be no agreed or enforced set of values, beyond tolerance: thus the political possibility of various types. It is as

though citizens are allowed to decide for themselves what will constitute their own goods. However, for Plato, this is not a serious psychological possibility: humans need a socially grounded culture to internalize.<sup>105</sup> A person may decide, say, to be a politician, but such a decision is superficial and eminently shakeable by external events. By historical luck the person may succeed at the appearance of state-craft, but Plato's point is that this is thin stuff. And so, even in democracy's finest hour, when it appears a many-colored fabric, full of different individuals each performing their own tasks, Plato's point is that this cannot be more than appearance. For although at that moment the citizens will not all be shifting their characters, they will all have characters which are *shiftable*. Thus their characters are unsteady, however firm they may appear.

Williams concludes:

There have been those who thought that the working classes were naturally of powerful and disorderly desires, and had to be kept in their place. There have been those who thought that they were goodhearted and loyal fellows of no great gifts who could recognize their natural superiors and, unless stirred up, keep themselves in their place. There can have been few who have thought both; Plato in the *Republic* comes close to being such a one.<sup>106</sup>

This thought is amusing, but not absurd. Indeed, if one takes the role of internalization seriously, it would seem to follow that in one political environment the working class will be a disorderly mob that has to be kept in its place, while in another it will consist in good-hearted fellows who recognize their superiors. Again and again, what presents itself as an absurdity dissolves once one takes seriously the idea that humans are dependent on internalization for acquiring psychological structure.

The initial appearance of absurdity depends once more on assuming that psychic parts are invulnerable to cultural influence. If the appetitive part must be in the conflictual relation with reason in which it is originally identified, then the working class will have to be a direct manifestation of contentious appetite. If intrapsychic conflict is unavoidable, then, given the analogy, so is political conflict. It will then look, just as it does to Williams, that when the obfuscating mask is pulled away we will see that Plato's just polis has the same repressive structure that Plato himself diagnoses in oligarchy. And, I think, it is tempting to go along with Williams' argument in part because Plato's ideal polis does look to us as though it has repressive features.<sup>107</sup> But the point of the present argument is not to rescue Plato's polis, it is to understand the psychological

basis of the isomorphism. Once one sees that the isomorphism is not a mere analogy, but is grounded in internalization and externalization, one sees that there is room to influence the shape and content of the psychic parts, and this allows room to influence the specific type of say, appetitive person, which in turn allows room to influence the specific type of appetitive class in the polis. This is hard to see in part because Plato concentrates so much on pathology, and pathological structures are inherently conflictual. Plato's psychology, like Freud's, is "wisdom won from illness."<sup>108</sup> Plato finds himself in a pathological social situation,<sup>109</sup> and, given his psychological principles, he deduces that this pathology is both cause and manifestation of pathology within the psyche. And it is his task to work his way back from the conceptualization of this pathology towards a conception of health.<sup>110</sup> His strategy was to assume a dynamic psychological relation between psyche and polis, and to construct an idealized genealogy of illness.<sup>111</sup>

For Plato, the hallmark of pathology is a lack of harmonious relations between inside and outside. That is one reason why the principles of internalization and externalization have been difficult to recognize. For it is a sign of oligarchy being a pathological structure that it cannot simply be internalized and externalized without further ado. The oligarchical father does externalize the structure of his psyche. And it is such externalizations which shape the oligarchical polis: by encouraging one class to accumulate wealth, the other class to forfeit it. The son, for his part, does internalize the polis influences. But because oligarchy is a pathological configuration, the internalization cannot stably reproduce the psychic structure of the previous generation. The instability is manifest in the inability of inside and outside to maintain a mirroring relation—and in the ensuing failure of the son to grow up in the image of the father. All this in spite of the fact that internalization and externalization are basic psychological activities.<sup>112</sup>

The point of Plato's argument is to show that there is only one relatively stable equilibrium position between inside and outside.<sup>113</sup> Only the just polis and its citizens are so structured that the various internalizations and externalizations will maintain harmony in each and harmony between them. Justice, for Plato, is a certain harmony within the psyche; it is also a certain harmony within the polis.<sup>114</sup> But now we can see that each of these harmonies is possible only if there is a larger harmony—between inside and outside—which encompasses and explains them.<sup>115</sup> Justice when properly understood is each part, inside and outside, doing its own task. That is why it is ultimately misleading to think

of there being merely an analogy between polis and psyche.<sup>116</sup> That is how it might look at the beginning of inquiry, but not how it should look at the end. When it is first introduced, the isomorphism may appear to be little more than an argumentative device. But then we, at that stage, are deep in the cave, confronted by what appear to be contradictory arguments about whether justice is good or bad.<sup>117</sup> The remainder of the *Republic* works through these contradictions, and what we come to see is that, roughly speaking, psyche is internalized polis and polis is externalized psyche. What initially appeared as two things which stood in a merely analogous relation come to appear as the internal and external workings of a psychological universe which may exist in various states of harmony or disharmony.

## 5. Poetic Justice

Internalization and externalization also explain why, for Plato, poetry corrupts our psyches. Given our psychology, there are two features of poetry which make it an especially potent drug. First, the music and rhythms with which poetry is expressed pour directly into our psyches.<sup>118</sup> Second, poetry tends to be expressed in imitative style: the characters speak as though from their own first-personal perspectives.<sup>119</sup> In this way, poetry can preserve the first-personal perspective throughout its transmissions.<sup>120</sup> Whether we are poet, performer or audience, we imaginatively take up the perspective of the characters: even the best of us abandon ourselves and imaginatively take up their feelings.<sup>121</sup> It is as though imitation blurs the boundary between inside and outside. Through imitation we get outside ourselves imaginatively, but psychologically we take the outside in. By pretending to be these characters, we unconsciously shape our characters around them.<sup>122</sup> The mimetic poet, says Plato, sets up a bad constitution in the psyche of each person (*τὸν μιμητικὸν ποιητὴν φήσομεν κακὴν πολιτείαν ἵδιᾳ ἐκάστου τῇ ψυχῇ ἐμποιεῖν*).<sup>123</sup>

Poetry feeds our psychological hunger to take things in, but it feeds us a diet of fantasy.<sup>124</sup> Its ability to draw us into a world of illusion indicates that it is appealing to a primitive level of mental functioning: Plato calls it “a vulgar part” (*τῶν φαύλων τὸ*) of the psyche.<sup>125</sup> For Plato, poetry has a hotline to the appetites.<sup>126</sup> It is able to bypass reason, the faculty which corrects for false appearance,<sup>127</sup> and go straight to the psychic gut. So while reason may tell us to be moderate in our grief, poetry en-

courages lamentation, excess and loss of control.<sup>128</sup> Poetry thus sets us up for intrapsychic conflict.<sup>129</sup> For poetry encourages “the irrational part” (*τὸ ἀλόγιστον*) of us to hold on to fantasy in spite of reason’s corrections. It establishes a split-off part of the psyche to which reason is not accessible. And that is why poetry cannot, for Plato, be just a stage in the developmental cave we work our way through. Other images may generate conflicts that lead us towards reality,<sup>130</sup> but poetic imitations keep us imprisoned at that level. So, on the one hand, poetry promotes intrapsychic conflict; on the other, it keeps us unconscious of that conflict, for the irrational part of our psyche cannot hear reason’s corrections. That is why poetry, with its throbbing rhythms and beating of breasts, appeals equally to the nondescript mob in the theatre and to the best among us.<sup>131</sup>

But if poetry goes straight to the lower part of the psyche, that is where it must come from. First, imitation by its very nature encourages poet, actor, and audience to go through the same motions. Although imitation is only play,<sup>132</sup> it is in this play that we unconsciously shape our psyches.<sup>133</sup> If poetic imitation sets our appetites in motion, it is reasonable to infer similar motions within the poet. Second, when a part of our psyche is strengthened from outside, it tends to be by an interpsychic manifestation of that very same part of the psyche. So, for example, the budding democrat’s appetites are reinforced by the appetitive thugs on the block.<sup>134</sup> The fact that poetry deals in fantasy and the throbbing lamentations of the irrational part of the psyche testifies to its lineage. Third, when Plato in his thought experiment wants to move from a minimal polis to a fevered one, he adds imitators (*οἱ μιμηταί*): poets, actors, rhapsodes, chorus dancers, theatrical managers.<sup>135</sup> He takes himself to be introducing a pathogen into a healthy organism. And the disease the polis contracts is *pleonexia*: the polis gives itself over to the unlimited acquisition of wealth.<sup>136</sup> Only after the polis is rid of poets who tell tales of gods eating, fighting, and deceiving each other, does Plato conclude that he has purged the fevered polis.<sup>137</sup> Introduce the poets and the polis becomes pleonectic, banish them and you cure it. Finally, as we have seen, *logos* follows and fits the character of the psyche.<sup>138</sup> If poetry is an appetitive falsehood, it must come from an appetitive affection in the psyche. And so it seems that just as law in a good society is an externalization of reason (of the philosopher king [who has already internalized the eternal realities]), so poetry seems to be an externalization of the irrational part [of the poet (who may already have internalized appetitive-poetic elements of culture)].<sup>139</sup>

We can see these appetites in the gods. The gods of the poets spend

their time castrating and devouring each other, they are constantly at war, and tend to engage in singleminded pursuit of satisfaction.<sup>140</sup> In short, these gods behave like lawless, unnecessary appetites;<sup>141</sup> and, given Plato's psychology, it seems reasonable to hypothesize that this is just what they are: appetites externalized in Olympus. A moment's reflection will show that there is nowhere else for them to go. Plato calls the lawless appetites "something wild and terrible" ( $\delta\epsilon\nu\delta\omega\tau\iota\ \kappa\alpha\iota\ \ddot{\alpha}\gamma\rho\tau\iota\omega\tau$ ) within us.<sup>142</sup> He speaks of Eros as a "tyrant within" ( $\tau\upmu\rho\alpha\nu\nu\omega\tau\ \epsilon\nu\delta\omega\tau$ ) the psyche.<sup>143</sup> Undisciplined appetites are all powerful within, so when they are externalized it makes sense that they should be represented as tremendously powerful. They need a virtually transcendent arena in which to struggle.<sup>144</sup> And so externalization from inside the poet's psyche turns out also to be an inversion: from bottom of the psyche to top of the world. These poetic myths provide a cultural template for youths to internalize, thus inverting their own psyches and, inevitably, the societies in which they live. Children, says Plato, will come to think there is nothing wrong in punishing their father to the limit, in fighting with their family and fellow citizens, if they think they are only following in the gods' footsteps.<sup>145</sup> And it is precisely by those acts, Plato thinks, that the tyrant is born.<sup>146</sup> According to legend, a person who eats human entrails is turned into a wolf; just so, the person who sheds the blood of the tribe by unjust accusations against fellow citizens, who banishes and slays them, has "tasted kindred blood," and is transformed into a tyrant. The tyrant is formed by transgressing the basic norms of human relations. In fact, the tyrant is behaving towards other humans as the Homeric gods behave towards each other. Plato criticizes Euripides for praising tyranny as "godlike" ( $\iota\sigma\omega\theta\epsilon\omega\tau$ ); but he is objecting not so much to the description, as to the fact that it is being used as a form of praise.<sup>147</sup> Tyranny is an imitation of the Homeric divine: but there is nothing praiseworthy about that.

This brings us to the most serious charge against the poets: they provide not only an externalization of the appetites, they also provide a *legitimation* of them. That is why the poetic myths are the "greatest lie about the greatest things," "an ugly falsehood."<sup>148</sup> The poets externalize their appetites, but their poetry sends them upwards as well as outwards. When the appetitive gods are reinternalized, it is now with a normative tinge.<sup>149</sup> Since the young are not able to distinguish myth from reality,<sup>150</sup> the tales they hear at their mother's knee provide the means by which the appetites can travel up and infect the norms and values of the developing person. In youth, we begin taking in psychological content and structure,

before we know how to distinguish truth from falsity. At a later stage of development, we attempt to take in true beliefs and expel falsehoods.<sup>151</sup> However, if we already have a falsehood inside our psyches, even in mythic form, we will end up taking in more and more falsehood (as though it were true) and getting rid of more and more truth (as though it were false). Introduce this initial virus, and our intake-expulsion machine will start pumping in the wrong direction.<sup>152</sup> That is why having falsehood inside the psyche is what humans loathe most of all.<sup>153</sup> And that, for Plato, is what mimetic poetry introduces: a falsehood taken as truth ( $\omega\zeta\ \ddot{\alpha}\lambda\eta\theta\omega\zeta\ \psi\epsilon\delta\omega\zeta$ ).<sup>154</sup>

Plato is charting the interpsychic and intrapsychic vicissitudes of the appetites. He is following the fate of the poetic trajectory: and what he finds is that the externalized appetites will tend to return, strengthened and legitimated. Poetry thus provides both a legitimization of the appetites and a cultural template for tyranny. One can see this in Plato's account of the rise of the tyrant. The tyrant is a child of democracy: the son of democratic man.<sup>155</sup> The democratic father is himself a compromise formation, shaped by a thrifty, oligarchical father, who encouraged only the acquisitive appetites, and a 'sophisticated' element which encouraged the unnecessary appetites. The pathology of this solution is again revealed by the instability between inside and outside. The son is brought up in the ways of the father, but is thereby susceptible to lawless influence from outside. It is the "dread magicians" ( $\omega\iota\ \delta\epsilon\nu\omega\iota\ \mu\alpha\gamma\omega\tau$ ) who both whet his lawless appetites and encourage him to expel from his psyche any remnants of shame which would keep the appetites in check. That the intake-expulsion machine is pumping in the wrong direction is testimony to there being a falsehood taken as truth within. And just as, *intrapsychically*, the lawless appetites overtook the original, better ones in his psyche, so, *interpsychically*, the tyrant comes to feel *justified* ( $\ddot{\alpha}\xi\iota\omega\omega$ ) in taking over his parent's estate; and then going on to rob, punish, and enslave family, friends, and fellow citizens.<sup>156</sup> In fact, the tyrant reenacts on the interpsychic stage of the polis the situation that exists inside his psyche: he must expel or get rid of the brave, wise, and wealthy, treating them as his enemy.<sup>157</sup> "A fine purgation," Plato says, "and just the opposite of what a physician does with our bodies: for while they remove the worst and leave the best, the tyrant does the opposite."<sup>158</sup> He recreates the polis in the image of his psyche.

And the poet gives him the cultural vehicle by which he can, at least to his own satisfaction, legitimize his acts. Hearing tales about the warring gods, Plato says, children will be encouraged to think this type of

behavior appropriate.<sup>159</sup> The gods of the poets are the lawless appetites externalized in Olympus: the tyrant brings this lawlessness back to the polis—sometimes literally. The tyrant is often someone who, because of previous attacks on society, has been banished from the polis.<sup>160</sup> There he remains poised for a triumphant return in the name of democracy; which for Plato is nothing more than a lawless society of appetites. Plato's point is that if you really want to get rid of the tyrant, you also have to get rid of the cultural vehicles that make him look attractive: you must also banish his poetical counterpart. For it is the poets who "draw the constitutions towards tyrannies and democracies."<sup>161</sup>

One might say that the tyrant acts out what the poet only dreams; but, for Plato, both the poet and the tyrant are dreamers, though in slightly different ways. To understand the tyrant, Plato says, we must not settle for his outward appearance, the external pomp and circumstance—we must even strip him of the garb in which tragedy clothes him<sup>162</sup>—and must, in thought, enter into his character.<sup>163</sup> What we find inside is a tyranny of lawless desires.<sup>164</sup> These are the desires we encounter in our dreams, when the rational part of our psyche sleeps, and the wild and animal part wakes up. "There is nothing it will not dare to do at that time, free of any control by shame or prudence. It does not hesitate to attempt sexual intercourse with a mother or anyone else—man, god, or beast; it will commit any foul murder and does not refrain from any kind of food. In a word it will not fall short of any folly or shameless deed."<sup>165</sup> These, of course, are the very deeds with which the gods of the poets occupy themselves. Indeed, the tyrant is a parricide;<sup>166</sup> and parricide is the founding act of Homeric heaven. It is this dreamworld that the poets have externalized in Olympus, and which the tyrant has reinternalized. The dreams of the poets enable the tyrant to turn his waking life into a bad dream: a nightmare.

Poet and tyrant are essentially dreaming the same dream; indeed, they are bedfellows. From Plato's perspective, poet and tyrant are the same type of person: a Dr. Jekyll and Mr. Hyde of the appetites. Both have organized themselves around their appetites, though they have different strategies for dealing with them. The tyrant keeps his appetites inside: because of them he outwardly enslaves the polis and inwardly is enslaved by them.<sup>167</sup> The poet externalizes his appetites: but there they form a cultural template which, when reinternalized, enslaves us all. Poet and tyrant ultimately enslave us, but while the tyrant enforces external compliance, poetic enslavement reaches inside the psyche and reorganizes it so that we remain unconscious of our slavery.

That is why the poets must be banished from the polis. One might say that Plato is recapitulating the poet's activity, only at a philosophical level. The poets, after all, have externalized their appetites, setting them up outside the polis in a heavenly beyond. What Plato sees is that the 'poetic solution' to the problem of the appetites in fact provokes a psychosocial disaster. The Platonic solution is inspired by his psychological principles. Plato knows that every externalization is fodder for internalization; and his 'final solution' is designed to put an end to this cycle. The important thing for Plato is not to get the poets out of the polis, so much as to get the appetites out of culture. This, he thinks, can be accomplished only by banishing the poets.<sup>168</sup>

Of course, there is plenty of room to doubt whether Plato's solution is called for or whether it would be successful. Does poetry not serve a healthy function? Is poetry not more (or other) than mimesis? Would the banished not find another way to return, if not from poetic heaven, then from beyond the philosophical pale? Is Plato's prescription so removed from anything we have ever experienced that we have no idea what is being prescribed? Rather than try to answer these questions, I shall close by explaining why we have only recently become ready to evaluate Plato's argument from a psychological perspective. Most recent discussions of the psychological value of art rely on an early psychoanalytic conception of the mind. The mind, on this conception, is divided along the lines of repression. The point of therapy was to loosen repression so that the unconscious could express itself, if only in words. In this context the creation and enjoyment of art appeared as another socially acceptable way of expressing unconscious forces. Thus artistic creation and appreciation came to be seen as therapeutic. As psychoanalytic theory developed, it became less concerned with the unconscious *per se*, and more concerned with the structure of the psyche. The psychoanalytic valuation of art has not kept pace with the development of theory.<sup>169</sup> In fact, Plato's remains one of the few discussions of the psychological value of art within the context of a structural theory of the psyche. Plato's point is that it is not enough to assume that the release of the repressed is a good thing. If one wants to justify art from a psychological perspective, one must understand its role within the context of a structured psyche. And that may require an account of the psychological transactions inside and outside the psyche. This is a challenge which, it seems to me, we are only now ready to take up once again.<sup>170</sup>

## Notes

1. There is even linguistic pressure on us to make that assumption. For although “psychology” is an English word, it comes almost directly from Greek; and while the English word carries the broad connotation of the science of mental activity, its Greek counterpart would be an account or logos of the psyche.

2. In the parlance of contemporary psychoanalysis, it leaves out Plato’s object-relations theory. Indeed, it leaves out the possibility of object-relations theory being an element of Plato’s psychology. Freud, of course, understood that a person’s ego and superego were formed around internalizations of parental figures. In the analytic situation, he concentrated on the intrapsychic configurations of the analysand, but he recognized that these configurations were due in part to interpsychic relations. See e.g., Freud, “Mourning and Melancholia,” *The Complete Psychological Works of Sigmund Freud* (London: Hogarth Press, 1957-81) XIV: 249-50; *The Ego and the Id*, XIX: 29-31; and my *Love and Its Place in Nature* (New York: Farrar, Straus and Giroux, 1990; London: Faber and Faber, 1992) chapter 6. For an introduction to post-Freudian object-relations theory see, e.g., Melanie Klein *Love, Guilt and Reparation*, and *Envy and Gratitude* (London: Hogarth Press, 1981, 1984); D.W. Winnicott, *Through Paediatrics to Psycho-Analysis* and *The Maturational Process and the Facilitating Environment* (London: Hogarth Press, 1975, 1982); W. R. D. Fairbairn, *Psycho-analytic Studies of the Personality* (London: Routledge & Kegan Paul, 1984); Margaret Mahler, *On Human Symbiosis and the Vicissitudes of Individuation* (New York: International Universities Press, 1967); Otto Kernberg, *Internal World and External Reality* (New York: Aronson, 1980).

3. See, e.g., Iris Murdoch, *The Fire and the Sun: Why Plato Banished the Artists* (Oxford: Clarendon, 1977); G. R. F. Ferrari, “Plato and Poetry,” *The Cambridge History of Literary Criticism*, vol. I, ed. G. Kennedy (Cambridge: Cambridge University Press, 1989); D. Halliwell, *Plato: Republic 10* (Wiltshire: Aris & Philips, 1988); Alexander Nehamas, “Plato on Imitation and Poetry in *Republic 10*,” in *Plato on Beauty, Wisdom and the Arts*, ed. J. Moravcsik and P. Temko (Totowa, N.J.: Rowman & Littlefield, 1982, 47-78).

4. II.358B, cp. 366E.

5. II.368D-E; IV.434D.

6. II.377A-B: my emphasis.

7. II.377B.

8. II.378D-E. cp. V.449D where Plato says that the constitution of the community of women and children makes all the difference to the constitution of the state.

9. II.380B-C.

10. See, e.g., III.386A, IV.424E.

11. II.381E.

12. III.387.

13. III.395C.

14. III.411E-412A; IV.424C-D.

15. V.466E-467A.

16. IV.420D-499D; cp. 430A-C; 441E-442C; V.465D-466D.

17. II.395D.

18. Cp. II.366C-D; IV.424E; VI.492A-493A; 495A-B; 496C; 499B-C.

19. IV.422E.

20. IV.422E-423D.

21. See Books VIII-IX.

22. VIII.551D.

23. VIII.553C-554E.

24. VIII.553C-D, cp. IX.581C.

25. VIII.554A.

26. VIII.554C-D. I do not believe that Plato’s conception of “forcible restraint” should be equated with Freud’s concept of repression, though there are of course similarities. For Freud, repression is itself unconscious, it is dynamically motivated, and the repressed is unconscious but continues to exercise influence in hidden ways. Plato does not suggest that the “forcibly restrained” is thereby rendered unconscious, or that these intrapsychic struggles must, by nature of the process, occur unconsciously.

27. VIII.554D.

28. Similarly, just as the democratic polis lets a hundred flowers bloom (VIII.557B-C), so democratic man is ‘manifold’ ( $\pi\alpha\tau\omega\delta\alpha\pi\omega\zeta$ : 561E). The democratic man is a compromise formation: an attempted solution to the conflicting demands of reason and appetite (550A-B). However, that the compromise fails is testified to by the emergence of oligarchic man in the next generation. The tyrant is just a mess. (I shall discuss the democrat and tyrant below.) For Plato, a human being, looked at from the outside, is only apparently a unity (IX.588C-E); whether each forms a genuine unity depends on the integration of the (potentially) disparate bits of the psyche.

29. Plato says (IV.433D) that in the just polis each person, in performing the task which suits his or her nature, will be not a multiplicity, but a unity. (See also IV.443E.) This suggests that a healthy polis encourages the development of healthy psyches: people who achieve the degree of psychic unity of which their character-types are capable. Injustice, by contrast, is a kind of civil war both in polis and psyche (IV.444A-B).

30. III.395C-D.

31. III.401D-E.

32. III.401 B-C. By contrast. cp. 401 C-D.

33. IV.424C.

34. IV.424D.

35. IX.590E.

36. I should stress that I here stipulate “internalization” to mean the process, whatever it is, that Plato thought grounded cultural influence. For an introduc-

tion to the concept of internalization as it occurs in the modern psychoanalytic tradition, see, e.g., Roy Schafer, *Aspects of Internalization* (Madison, Conn.: International Universities Press, 1990); Hans Loewald, "On Internalization" and "Internalization, Separation, Mourning and the Superego," in *Papers on Psychoanalysis* (New Haven, Conn.: Yale University Press, 1980); J. LaPlanche and J-B Pontalis, *The Language of Psychoanalysis* (London: Hogarth Press, 1983), 205-208, 211-12, 226-27, 229-31; B. Moore and B. Fine, eds., *Psychoanalytic Terms and Concepts* (New Haven, Conn.: Yale University Press, 1990), 102-103, 109-10; R. D. Hinshelwood, *A Dictionary of Kleinian Thought* (London: Free Association, 1991), 68-83, 319-21, 330-34; Arnold Model, *Object Love and Reality* (Madison, Conn.: International Universities Press, 1985).

37. VIII.544D-E. See also IV.435E, quoted below.

38. III.402D. Cp. VI.535B; III.401A; IX.577A.

39. IV.435C.

40. V.435E-436A; see also 441C.

41. I shall discuss pathological forms of polis in section 4.

42. VI.500B-C.

43. VI.500E.

44. VI.500C. Cp. 484C: They have a paradigm of the reality of things in their psyche. (See also 490B). This is the step which Charles Taylor omits from his account of Plato in *Sources of the Self* (Cambridge, Mass.: Harvard University Press, 1989), 121-26.

45. VI.500C-D.

46. VI.484C, 490B. Some such internalization is necessary, Plato thinks, for a person with a philosophical nature to achieve excellence. (VI.492A) This may be through a proper education, but with poor upbringing even a philosophical nature is destroyed and corrupted. (VI.495A-B) Such a person is then capable of the greatest evils, and his only hope is divine inspiration. That is why a person of philosophical nature ought to shun political life in a pathological polis: he must take care of the "constitution inside himself" ( $\tau\grave{\eta}\nu\ \grave{\epsilon}\nu\ \alpha\acute{\nu}\tau\acute{\omega}\ \pi\o\lambda\iota\tau\acute{\epsilon}\iota\alpha\acute{\nu}$ ) and not allow cultural influences to "undo the state of his psyche."

47. IX.592B.

48. Socrates argues that education is not, as the sophists think, a matter of putting knowledge into a psyche, but rather more like turning the eye from the dark (world of becoming) to the light (world of realities). (VII.518B-E) This metaphor may have impeded understanding of Plato's psychology. For Plato is not here saying that internalization does not take place in education, he is rather explaining how internalization comes about. It is more, he thinks, than learning a few sophistical speeches. The point of turning one's gaze towards reality is not just to gawp at it like a bewildered tourist; it is to take reality in, be educated by it.

49. III.400D-E.

50. III.400E-401A.

51. II.376E-377A.

52. II.377A.

53. II.382B-C.

54. II.382C-D.

55. II.382B.

56. I shall discuss the intra- and interpsychic transformations in the following sections.

57. VI.485D. (For the economic model in psychoanalysis, see e.g., Freud, "The Unconscious," XIV.181; *Studies on Hysteria*, II.17, 86, 166-67; "The Neuro-Psychoses of Defense," III.48-49).

58. II.368D-369A.

59. In E. N. Lee, A. P. D. Mourelatos, and R. M. Rorty, eds., *Exegesis and Argument: Studies in Greek Philosophy Presented to Gregory Vlastos, Phronesis: Supplementary Volume I* (Assen, The Netherlands: Van Gorcum, 1973). This essay has influenced a generation of philosophers, myself included. I turn to it here because I have come to believe, first, that the argument is unsuccessful; second, that in coming to understand why it is unsuccessful we will better understand our own tendency to misread Plato's psychology.

60. Williams derives this from 435E. See "Analogy," 196-97.

61. Derived from 435A-B.

62. "Analogy," 197.

63. See below.

64. IV.435A-B.

65. IV.435E.

66. Terence Irwin argues that the whole-part rule does not play a role in the argument of Book IV, and focuses instead on Macrocosm-Microcosm rule (MM): the structure of the state is analogous to the structure of the psyche. (*Plato's Moral Theory: The Early and Middle Dialogues* [Oxford: Clarendon, 1977], 331 n. 29). The MM is true, but it does not give us the psychological principles which ensure its truth. John Cooper also provides criticism of the whole-part rule in "The Psychology of Justice in Plato," *American Philosophical Quarterly* 14 (1977), n. 7.

67. "Analogy," 198, cp. "the ineliminable tension" on 200 ff.

68. This is the strongest version of the whole-part rule we are legitimately entitled to attribute to Plato.

69. Williams comes close to accepting this when he later adopts the "predominant section rule" which I shall discuss below.

70. "Analogy," 199.

71. Since we have substituted (a'') for (a'), there is no longer reason to believe that everyone in a just polis is just. We therefore look at the psyche of an appetitive-type person.

72. VIII.558D-559C.

73. VIII.554B-E.

74. IV.421C.

75. Cp. IX.576C-D.

76. Essentially the same problem occurs in Freud's discussion of the id. Freud often describes the id as not listening to reason. But he is so describing it in the context of trying to make clear the dynamic structure of neurotic pathology. There is another conception of the id, manifest in his dictum "Where it was there I shall become," which allows the possibility of the appetites harkening to reason. See my *Love and Its Place in Nature*, chapter 6.

77. Of course, there is truth in the claim that money is a means to satisfy bodily appetites (IX.580E-581A), but that is not the whole truth. The oligarchic man, for example, is not using money just as a means to satisfy his bodily appetites: indeed, he keeps these appetites under control precisely because he has developed an appetite for money and property (VIII.553C-554C).

78. IX.58OD-581E; cp. VIII.553C-554E.

79. In fact the variation can be much more fine-grained than I have indicated. See the explication by C. D. C. Reeve in *Philosopher-Kings* (Princeton: Princeton University Press, 1988, 5-9, 135-53), a book which I found inspiring and to which I am indebted. I am here both trying to use that account and to show how much it depends on internalization as a basic psychological activity. See also John Cooper in "Plato's Theory of Human Motivation," *History of Philosophy Quarterly* 1, 1984; and Richard Kraut's account of normative rule in "Reason and Justice in Plato's *Republic*," *Exegesis and Argument*, 207-24.

80. X.581.

81. IV.43ID-E.

82. IX.58IC-D.

83. IV.43ID-E, 433C-D.

84. IV.433E-434A. See Reeve, *Philosopher-Kings*, 246-47.

85. IV.431E-432A.

86. IX.58IC-D.

87. IV.429B-430A.

88. IV.429C.

89. Plato, as we have seen, believes a change in musical modes will ultimately upset constitutional laws: it is precisely because lawlessness is internalized with the music that it is subsequently externalized in attacks on established business relations, on the laws and the constitution (IV.424C-E). I shall discuss this further in section 4.

90. VIII.558-562.

91. VIII.554B-555B.

92. VIII.554D.

93. Roughly speaking because, as we have seen, Plato's conception of forcible restraint is not identical to Freud's conception of repression. Yet Plato believes of forcible restraint, as Freud believed of repression, that it is an ultimately unsuccessful means of warding off unwanted desires.

94. VIII.556B-C.

95. VIII.559D.

96. VIII.555C-E.

97. VIII.557A.

98. VIII.559E.

99. VIII.560. Plato's account of faction vs. counterfaction struggling within the psyche bears some similarity to Freud's account of cathexis and counter-cathexis in a neurotic struggle—although there is no evidence that Plato thought this intrapsychic and interpsychic struggle was unconscious. (Cp. e.g., "Further Remarks on the Neuro-Psychoses of Defense," III:169-70).

100. VIII.560B-C.

101. VIII.557C ff. Quoted by Williams, "Analogy," 201.

102. VIII.561D ff.

103. "Analogy," 201. Note that by now Williams has put in place of the whole-part rule, another which he calls the *predominant section rule*: A city is F if and only if the leading, most influential or predominant citizens are F.

104. VIII.557B-558A.

105. Plato does make an exception for those who have been divinely inspired: e.g., VI.492A, 496C-497A, 499B-C.

106. "Analogy," 204.

107. I suspect that Williams' formal objections to the analogy are fueled by a democrat's suspicion of Plato's conservative political theory: in particular by what he takes to be an ultimately repressive relation between ruling class and ruled. From a democratic perspective the means and organization of society ought to be transparent to all, while Plato advocates feeding the appetitive class a diet of noble falsehoods. It is, of course, beyond the scope of this paper to respond to this type of objection. But I would like to note in passing: (1) Such an objection does not itself constitute an objection to the idea of an isomorphism between polis and psyche. (2) Plato himself issues a challenge to the idea of transparency. One of the motivations for the 'noble falsehood' is that one cannot just assume that, say, the freedom of information act guarantees freedom of information: one has to take into account what the subjective meaning of this (external) information will be. And once one does so, Plato thinks, one can only get this information across if one presents it in certain mythic forms which, strictly speaking, are not true. Each side will think the other is restricting information, one because of the alteration in form, the other because the idea of subjective understanding is being ignored.

108. The phrase is from Thomas Mann's description of psychoanalysis. See "Freud and the Future," in *Freud, Goethe, Wagner* (New York: Knopf, 1937).

109. See, e.g., VI.497A-B, 496C-D, 488A-489B, 499B-C.

110. Virtue or excellence, Plato says, is a certain sort of health (IV.444D).

111. In a sense Plato has again to recapitulate the poet, only at a philosophical level. Socrates must tell a tale in which the just man is stripped of all the outer trappings and the glittering prizes—which from a conventional perspective are all the rewards there are—are given to the unjust man (II.360E-361D). This, in effect, is what the poets have already done (362E-366A). They have shaped a culture which values only the appearance of justice. By showing that it is nev-

ertheless better to be just, Plato is doing more than showing that justice will triumph even in the worst possible dialectical circumstances. He is trying to show that it will triumph in (what he took to be) the actual situation. From Plato's perspective, his argument has to take this shape if it is to be persuasive, for the worst-case scenario is the way things are. Plato thus starts out with poetic appearance in order to work through to a (nonpoetical) conclusion which penetrates beyond surface appearances.

112. So for any pathological structure  $F^*$ , one should not expect that an  $F^*$  polis is an immediate and simple externalization of  $F^*$  citizens. Nor should one think that  $F^*$  citizens are shaped by a simple internalization of the structure of the  $F^*$  polis. The whole point of  $F^*$  being pathological is that no such simple mirroring relation can occur. So, for example, the democratic polis is shaped not only by the degenerate son of oligarchy, but also by the rebellious poor (556C-557A). However, the rebellious poor also had their psyches shaped via internalizations of previous externalizations of the oligarchical rulers. And both they and democratic man—the metaphorical and literal sons of oligarchy—help to shape the democratic polis via externalization of structures in their psyches.

113. And this is built up by what Plato calls a “circle of growth,” which seems to be the opposite of the tale of degeneration: “a sound nurture and education if kept up creates good natures in the state, and sound natures in turn receiving an education of this sort develop into better persons than their predecessors” IV.424A-B. Although, of course, Plato thinks that even the ideal polis is subject to eventual decay: VIII.546A-547A.

114. IV.441D-E. See e.g., V.462C-D, 463E, 464A.

115. And thus I think the psychological principles of internalization and externalization can help us to address a long-standing interpretive problem: why did Plato think there was a relation between justice as a condition of the psyche—psychic justice—and conventionally recognized justice? (See, e.g., Gregory Vlastos, “Justice and Happiness in Plato's *Republic*,” *Platonic Studies* (Princeton, N.J.: Princeton University Press, 1981); David Sachs, “A Fallacy in Plato's *Republic*,” *Philosophical Review* 72 (1963): 141-58; Terence Irwin, *Plato's Moral Theory* (205-206, 331). I address this problem in “Plato's Politics of Narcissism,” which I presented at the memorial conference for Gregory Vlastos in May 1992 and which will appear in a volume dedicated to his memory.

116. In fact, Plato never uses the word “analogy” ( $\alpha\gamma\alpha\lambda\omega\gamma\iota\alpha$ ) to describe the relation between polis and psyche, though he is sometimes translated as though he did. See, e.g., Paul Shorey's translation of II.368E in the Loeb edition of *The Republic* (Cambridge, Mass.: Harvard University Press, 1982).

117. The contradictory arguments of Book I bear a significant resemblance to the problems which Plato says are provocative of thought (VII.523-524).

118. III.411; cp. e.g., 395C-D, 401B-D, 413C.

119. III.393B-D.

120. See, e.g., Ferrari, “Plato and Poetry,” esp. §§1,4.

121. X.605.

122. III.395C-396E, 378D, 398A-B, 401B-402A; X.605-606.

123. X.605B-C.

124. X.598B: *phantasma*, cp. 599A. In fact, Plato says that imitation gives us a fantasy of a fantasy—a second-order fantasy, but this depends on his metaphysical conception of ordinary empirical objects themselves being removed from reality. There are, obviously, important metaphysical objections to tragic poetry and art, but in this paper I am restricting my focus to the primarily psychological objections.

125. X.603A, 605A-B.

126. X.603A-B; cp. 605A-B.

127. X.601B, 602D-603B.

128. X.604-605.

129. E.g., X.604B; cp. 603D.

130. See, e.g., VII.523.

131. X.604E-605D.

132. X.602B.

133. X.606B.

134. VII.559E-560. See above, 184-85.

135. II.373B.

136. II.373D-E.

137. III.399E. The purgation is supposed to have occurred as an unconscious by-product of banishing the poets.

138. III.400D-E.

139. X.604D, 605B, 605E-606B.

140. II.377E-378D (Should the reader also be interested in the work of Melanie Klein, and wonder what she meant by “part objects,” the Homeric gods of which Plato complains are, I think, paradigms.)

141. IX.571B-D.

142. IX.572B.

143. IX.573D; cp. 573B.

144. Compare Freud on the omnipotence of archaic mind: e.g., *Totem and Taboo*, XIII:83-91, 186, 188; “The Uncanny,” XVII: 240-44.

145. II.378B-C.

146. IX.565D-566B.

147. VIII.568A-B.

148. II.377E.

149. Freud noticed that the superego often speaks with an iddih accent: it tends to take on a harsh, vindictive tone that testifies to some sort of commerce with the id. (See, e.g., *The Ego and the Id*, XIX: 36, 48-49, 52; “Neurosis and Psychosis,” XIX:151-52; *Inhibitions, Symptoms and Anxiety*, XX:115-16.) This was puzzling both because the superego's function is to help keep the id in place and because it is unclear how this commerce takes place. Plato in fact provides a satisfying explanation of how commerce between id and superego can occur via

## Akrasia and Agency in Plato's *Laws* and *Republic*

Christopher Bobonich

### 1. Introduction

Philosophers from Plato to Davidson have long found akratic action puzzling. One well-known sign of Plato's puzzlement is that although he rejects the possibility of akratic action in the *Protagoras*, he reworks his moral psychology in the *Republic* to accommodate its possibility. But what has not been noticed is that the *Republic* is not the final resolution of Plato's puzzlement. In the *Laws*, Plato returns to the problem of akratic action. As in the *Republic*, Plato accepts its possibility, but in the *Laws* he offers a new analysis of what goes on inside an agent who acts akratically. As the shift in Plato's position from the *Protagoras* to the *Republic* required a revision in Plato's moral psychology, so the new analysis of akratic action in the *Laws* requires an important revision in the moral psychology of the *Republic*. The *Republic*'s explanation of the nature of the conflict present in the agent who acts akratically makes essential reference to different parts of the soul which are themselves agentlike. The *Laws* explains akratic action without invoking agentlike parts of the soul, and the agentlike parts of the soul of the *Republic* are absent from the moral psychology of the *Laws*.

In section 2, I briefly sketch the theory of parts of the soul found in the *Republic* and the role that these parts play in its analysis of akratic action. I argue that the resulting theory of akratic action has serious problems and flaws. In section 3, I outline the new theory of akratic action found in the *Laws* and show how it differs from the

*Republic* theory and how it avoids the difficulties of that theory. In section 4, I offer support for my claim that the *Republic*'s agentlike parts of the soul are absent from the *Laws*. In the final section, I consider some of the connections between this change in Plato's moral psychology and his analysis of akratic action and issues in his moral and political philosophy.

## 2. Akrasia and Parts of the Soul in the *Republic*

A full explication of the *Republic*'s theory of parts of the soul would require a monograph of its own: here I shall sketch only the broad features that are most relevant to the topic at hand. The *Republic* divides the soul into three parts: the Reasoning part, the Spirited part, and the Desiring part (*τὸ λογιστικόν*, *τὸ θυμοειδές*, and *τὸ ἐπιθυμητικόν*).<sup>1</sup> Each of these three parts is agentlike: each is the subject of psychic states, activities, and capacities that are normally attributed to the whole person.<sup>2</sup> In particular, each part

1. has its own desires (*ἐπιθυμίας*), and can wish and want (*βούλεσθαι* and *ἐθέλειν*),
2. has its own pleasures,
3. has cognitive and linguistic capacities:
  - (i) each has beliefs,
  - (ii) each can communicate with the others: one part can persuade another and they can all agree, and
  - (iii) each can engage in some forms of reasoning.<sup>3</sup>

I share the view of many other recent commentators that we should not treat as merely metaphorical or read away these attributions of psychological states and activities to the parts of the soul.<sup>4</sup> For better or worse, Plato's moral psychology in the *Republic* is committed to the idea that every person is a compound of agentlike parts.<sup>5</sup> (I shall refer to the theory of parts of the soul found in the *Republic* as the 'Partition Theory'.)

The Partition Theory plays a fundamental role throughout the *Republic*. For example, in Book 4 the virtues are defined in terms of parts of the soul, and the fact that certain pleasures belong to certain parts of the soul plays a role in Plato's argument in Book 9 that the philosopher's life is the most pleasant possible life. What is most important for my present purposes is that the parts of the soul are essential to Plato's explanation of akratic action in the *Republic*. In

every instance of akratic action and every case of successful resistance to akratic action, an essential part of what is going on inside the agent is a conflict between at least two parts of the soul. In Book 4 of the *Republic*, Plato offers two ways of describing this conflict and its resolution. He thus offers two ways of describing what goes on inside an akratic agent. I, for reasons that will soon become evident, shall call one the 'Command Model,' and the other the 'Force Model'.<sup>6</sup>

We can see the origin of these two models at *Rep.* 437B1 ff. Plato has already stated the Principle of Opposites,

The same thing will not do or undergo opposites with respect to the same thing and in relation to the same thing at the same time [δῆλον δτι ταῦτον τάνατία ποιεῖν η̄ πάσχειν κατὰ ταῦτον γε καὶ πρὸς ταῦτον οὐκ ἐθελήσει ἄμα] (*Rep.* 436B8-9),

and disposed of a few alleged counterexamples to this principle.<sup>7</sup> He now goes on to sketch what are going to count as opposites in the case of the soul:

"Would you set down all such things as opposites to one another: assent to dissent, striving to take something to rejecting it, drawing something in to thrusting it away, whether they are actions or affections? That will not make any difference."

"None," he said, "but they are opposites."

"What about being thirsty and hungry and generally the desires [*τὰς ἐπιθυμίας*], and further, wishing and wanting? [*τὸ ἐθέλειν καὶ τὸ βούλεσθαι*]," I said. "Would you not set all these somewhere in the classes just mentioned? For example, will you not say that the soul of a man who desires either strives for what it desires or draws toward itself that which it wants to become its own; or again, that, insofar as the soul wishes that something be supplied to it, it nods assent to this itself as though someone posed a question and reaches out toward its attainment?"

"I shall."

"What of not-wanting, and not-wishing and not-desiring [*τὸ ἀβούλειν καὶ μὴ ἐθέλειν μηδὲ ἐπιθυμεῖν*], shall we not class these with the soul's thrusting away from itself and driving away from itself and generally with all the opposites of the former?"

"Of course." (*Rep.* 437B1-D1)<sup>8</sup>

Roughly put, Plato's overall strategy here is to show that there are cases in which people do opposites with respect to the same thing and in relation to the same thing at the same time. Applying the Principle of Opposites to these cases, it follows that these people are not single entities but are composite: there is a distinct subject for

each opposite. In order for this strategy to succeed, Plato needs to show that the conflicting desires, wishes, and so on that people have do count as opposites which fall under the Principle of Opposites. And in the above passage, Plato offers at least two ways of describing desiring, wanting, and wishing so as to make plausible their inclusion in the class of opposites. (I shall simply speak of desiring from now on.) At 437B1-4, he distinguishes assent and dissent and drawing in and pushing away as two distinct pairs of opposites. This first characterization of desiring as a form of interior assent allows Plato to re-describe the opposition between desiring and not-desiring in terms of the logical and psychological opposition between assent and dissent.<sup>9</sup> The second sort of description that Plato gives is in terms of opposed forces. The conflicting desires are described as forces acting in opposite directions. In particular, the possible object of desire is conceived of as either being drawn in or pushed away. As we shall see, Plato uses both these descriptions in his characterization of akratic action, and as we shall also see, there are difficulties about the relations between the two descriptions.

Let us thus examine Plato's star example, in this passage, of potentially akratic conflict—that of a thirsty man.

"So the soul of a man who is thirsty, insofar as it thirsts, wants nothing other than to drink, and reaches out for this and is impelled toward it."

"Obviously."

"If ever something draws it back when it is thirsting, would that not be something different in it from that which thirsts and leads it like a beast to drink? For of course, we say, the same thing would not do opposites concerning the same thing with the same part of itself at the same time."

"Certainly not."

"In the same way, it is not fair to say of the archer that his hands at the same time thrust the bow away and draw it near, but that one hand pushes it away and the other draws it in."

"Certainly."

"Would we assert that sometimes there are men who are thirsty but do not wish to drink?"

"Yes, many and often."

"What then should one say about them?" I said. "Is not there something in their soul ordering them to drink ( $\tauὸ\ \chiελεῦον$ ) and something forbidding ( $\tauὸ\ \chiωλῦον$ ) them to do so, something different that masters ( $\chiρωτοῦ$ ) that which orders?"

"I think so." (Rep. 439A9-C8)

Now with regard to the model of opposed forces, we see that it is here given a slightly different form. No longer is the object of desire the recipient of the force applied; rather the soul, or more precisely, one part of the soul, is seen as moving toward the object while the other part drags the soul back. Plato's idea here, in the Force Model, seems to be both similar and dissimilar in interesting ways to recent attempts to explain akratic action in terms of the motivational strength of the agent's desires. In particular, we might compare Plato's account to the following account of akratic action. Let me emphasize that the following is not Plato's account, nor should it be attributed to any particular contemporary philosopher—but it is, I think, a plausible beginning of an explanation of akratic action and it does resemble some recent analyses.

The account runs as follows. An agent believes that action X is her best option all things considered (ATC) and, in particular, that it is better ATC than action Y and believes that she can do either X or Y but not both. Nevertheless, the agent's desire to do Y is stronger than her desire to do X. Thus if the agent acts, she will do (or try to do) Y. And the simplest way to reach the conclusion that the agent will do Y is to postulate the generalization that an agent will do whatever she most strongly desires to do.<sup>10</sup> How close is this account to the one that Plato offers?

To begin, we should note that even if the account just given were the same as the one contained in the Force Model, the Force Model is only one of the two models of akratic action that Plato offers. Thus Plato's full account of akratic action is at least more complicated (and perhaps just more confused) than the one just sketched. Restricting ourselves now to just the Force Model, there are certain resemblances between it and the above non-Platonic account. We can distinguish, in the Force Model, on the one hand the ATC judgment that it is best not to drink and the corresponding desire not to drink and, on the other hand, the desire to drink. (Since what forbids comes to be from calculation and calculation is concerned with what is better and worse for the whole soul [442C5-8], we may take this as an ATC judgment of what is best. Plato does not explicitly say that a desire for what the agent believes to be ATC best is at work here, but he does hold [413A4-5 and 505D5-E4] that all people desire what they believe ATC best for themselves.) The two desires are tugs or pulls on the agent which can be stronger or weaker; thus Plato accepts the idea that desires can vary in strength.<sup>11</sup> And as the example of the thirsty man shows, the strength of the pull is distinct from and is not always directly proportional to the agent's judgment of the goodness of the object of the desire. (More precisely, the strength of the Desiring

part's desire is distinct from and is not always directly proportional to the Reasoning part's judgment of the goodness of the object of the desire. It also seems that the strength of the Reasoning part's desire is distinct from its own judgment of goodness. Below I shall take up the question of whether the Desiring part makes judgments of goodness.) Finally, we can fill out the Force Model for Plato by claiming that what the agent does is determined by the force of the opposing pulls: the agent will act in accordance with whatever desire pulls most strongly.<sup>12</sup>

But there remains an important difference between the two accounts. The non-Platonic explanation of akratic action sketched above invokes the idea of the strength of a desire, but it is a crucial part of its explanation of akratic action that a single agent can believe that X is better ATC than Y, but still desire more strongly to do Y. The Force Model, on the other hand, explains akratic action by positing that one agentlike part believes that X is better ATC than Y and thus desires to do X while a different agentlike part more strongly desires to do Y. It would be rash to claim that this feature of Plato's theory of akratic action renders it unacceptable since I have not shown, and obviously cannot try to do so here, that it really is possible for a single agent to believe that one course of action is best ATC and yet still desire more strongly to take another course of action. But it is natural to wonder whether we could not produce a theory of akratic action that is more satisfactory than the Force Model by keeping the idea that desires vary in strength in such a way that their ranking in terms of strength does not always correspond to the agent's ATC rankings of the objects of her desires, while dropping the notion of the parts of the soul.<sup>13</sup> It seems that the above distinction between an agent's evaluation and the strength of her desires suffices to explain akratic action without the further hypothesis of agentlike parts. Such an explanation will avoid the puzzles produced by the partitioning of the soul and do justice to the intuition that all we need to recognize in akratic action is a single agent, her judgments and desires. And this, I shall argue, is the theory that Plato adopts in the *Laws*.

Although the Force Model seems to invoke the notion of the strength of a desire, Plato does not investigate the notion further here. The strength of a desire is distinct from a judgment of the goodness of its object, but it is not clear what positive content Plato gives to the concept of strength. Is the strength of a desire, for example, identical with its felt intensity?<sup>14</sup> Nevertheless, it is worth emphasizing that insofar as Plato here characterizes desires in terms of drawing, pulling, and pushing and relies on the notion of the strength of these forces in order to explain why the agent acts one way rather

than another, he seems committed to the idea that psychic items such as desires interact in the same way that physical objects do, that is, by processes which we would describe as causal.<sup>15</sup> If this is right, we would like to know whether or not this commitment is new in the *Republic*. One reason for thinking it is new is the following. If one does not accept the possibility of akrasia, it may seem that one can explain agents' actions simply in terms of their beliefs and desires and the rational interactions among them. The attribution of causality to psychic items may thus seem unnecessary and may be rejected as superfluous or as somehow incompatible with the nature of psychic items (or the question may simply not arise).<sup>16</sup> Since akrasia involves a breakdown of rationality, we need some further explanation of the agent's action, and causality provides such an explanation. Although I cannot explore this issue further here, it raises questions which are central to a full understanding of Plato's philosophy of mind.

Nevertheless, as we have already noted, the Force Model is only one of Plato's models of akratic action. Starting at 439C2, Plato offers another description of the case of the thirsty man. The two opposed parts of the soul are now described as the one which commands and the one which forbids ( $\tauὸ\ \chiελεῦον$  and  $\tauὸ\ \chiωλῦον$  439C5-6). The parts, as conceived here, thus seem to issue imperatives, e.g., "Drink!" and "Don't drink!" But the question is: to whom are the imperatives addressed? We cannot say that they are addressed to the *whole person* if we conceive of the whole person as an entity above and beyond the parts of the soul: like Plato's city, his person is nothing more than a compound of its parts.

It is especially tempting to see the whole person as the 'agent of last resort' in the following sense: the whole person receives the conflicting imperatives, considers the reasons for each, makes an ATC judgment, and then acts or tries to bring about the action. Compare, for instance, the account Davidson attributes to Plato:

Here there are three actors on the stage: reason, desire, and the one who lets desire get the upper hand. The third actor is perhaps named 'the Will' (or 'Conscience'). It is up to the Will to decide who wins the battle. If the Will is strong, he gives the palm to reason; if he is weak he may allow pleasure or passion to get the upper hand.<sup>17</sup>

But this is not the sort of account we find here in Book 4 nor is it an attractive one. The Reasoning part has already made an ATC judgment of what is best for the whole person and we also face the problem, on this story, of explaining how the whole person could ever act akratically.

The most plausible suggestion is that the conflicting imperatives are addressed to the other parts. (E.g., 442B5-D1 and 589A6-B6. Each imperative may also be self-addressed, cf. 437C4-6.) The question now is how this conflict of imperatives is to be resolved. Unfortunately, Plato simply says here that one part “masters” the other (*κρατοῦν* 439C7, cf. 430E11-431D3). Plato does not tell us precisely how to understand this talk of one part mastering another, but there are two possibilities to be considered. First, it may be the case that what the agent does is simply determined, as it was in the Force Model, by the strength of the opposing pulls; that is, it may still be the case that the agent will act in accordance with her strongest desire. But if this is right, the interaction of the parts is limited to a competition between their respective desires in terms of strength. The conflicting imperatives seem to do no work in explaining why the agent acts one way rather than another.<sup>18</sup> But there is another possibility which gives more point to the imperatives, since it relies on the idea that the parts can communicate with one another. For example, Plato is committed to the idea that the Reasoning part can “persuade” the Desiring part. Here the transfer of information between the parts and the change in belief by one of the parts seem to be necessary to the explanation of why the agent acts one way rather than another. This latter possibility, however, raises some problems for Plato.<sup>19</sup>

If the Desiring part is persuaded, what is it persuaded of? Several passages (554C11-E5 and 442B5-D1)—one of which I shall discuss further in section 4—show that when the lower parts are persuaded what they are persuaded of is that it is best for them to obey. They are persuaded of this by some form of argument (through *λόγοι* 554D2) and thus must be capable of at least some simple form of reasoning: they must be capable of following the argument. In particular, what seems to be required is that the lower parts of the soul be able to judge of what is ATC best for them even if they cannot judge of what is ATC best for the whole soul (442C5-8).<sup>20</sup> But now the Desiring part seems to have both the original desire which conflicted with the judgment of the Reasoning part and an ATC judgment that it is best to go along with the Reasoning part. Hence there is the possibility of conflict within the Desiring part, and this raises serious difficulties for Plato.

The possibility of akrasia now recurs at the level of the Desiring part. The Desiring part has an ATC judgment about what is best for itself and also may have desires for some incompatible option. Plato has now merely pushed the problem of akrasia down to the level of the Desiring part and has no clear way of analyzing and resolving it.

Does Plato think that once the Desiring part makes an ATC judgment of what is best for itself, it will lose all of its desires for what is not ATC best or that at least the strength of these desires will be diminished? If so, what guarantees that persuading the Desiring part will have this effect? According to the description in Book 4 (437D2-439B1) of at least some of the desires of the Desiring part, such as those for food and drink, these desires are not directed at the good of the whole individual, nor at the good of the Desiring part, nor are they derivative from such desires. While this account of the desires for food and drink does not by itself rule out the possibility that these desires disappear or diminish in strength once an appropriate judgment is made by the Desiring part, it does make it difficult to see why this should happen.

There are other possibilities to consider here, but none seems attractive. These include: (1) the Desiring part's desire for its own good is always its strongest desire, (2) the Desiring part avoids akratic conflict within itself by preventing the formation of desires which prompt akratic action, and (3) the Desiring part can act in accordance with its own desire for what is ATC best for itself even if this is not its strongest desire. (1) merely seems to be *ad hoc*. (2) still leaves the puzzle of how the Desiring part manages such prevention given the nature of the contrary desires, and Plato's description of the persuasion of the Desiring part shows that this persuasion goes on while the Desiring part still has akratic desires (554C11-E5 and 589A6-B6).<sup>21</sup> Finally, (3) involves an exception to the idea that the strongest desire always wins, and if the Desiring part is capable of overriding its strongest desire, why not just assign a similar capacity to the Reasoning part in the first place?

But even if we could accept the idea that the Desiring part, once it is persuaded acts, for whatever reason, in accordance with its judgment of what is ATC best in the face of the incompatible desires, e.g., for food and drink, a deep problem remains for Plato. The very possibility of a conflict within the Desiring part between its desire for what is best for itself and its desires for food and drink calls its unity into question. Does not such a conflict within the Desiring part force us to apply the Principle of Opposites to the Desiring part itself and thus subdivide it further? Even without a full analysis of the nature of the opposition that warrants the application of the Principle of Opposites, it seems that if a conflict between a desire for what is ATC best for the whole soul and the sort of desire for drink described at 439A9 ff. cannot be found within a single subject, then a conflict between a desire for what is ATC best for the Desiring part and the same sort of

desire for drink cannot be found within a single subject.<sup>22</sup> If so, Plato will wind up with more parts of the soul than his theory allows for.

Could the Reasoning part ever be persuaded on this model? If not, how do the other parts ever get their way? Plato might hold that if neither part is persuaded, the agent will act in accordance with her strongest desire. This suggestion explains the agent's actions by reference to the strength of her desires and is subject to the same sorts of concerns as the Force Model: does the partitioning do any explanatory work, or can we give a satisfactory explanation which uses the idea of the strength of a desire but does not partition the soul? The persuasion of the Reasoning part, however, raises interesting questions since it is hard to see what epistemic or prudential reason the Reasoning part could have for changing its belief in response to the communications of the Desiring part. Thus if the Desiring part induces a change in the beliefs of the Reasoning part about what is ATC best, this will be an irrational change, and Plato seems to recognize such a possibility.<sup>23</sup> In this case, the agent does not act contrary to the ATC judgment of what is best that her Reasoning part has at the time of the action.

Finally, attributing to the Desiring part the capacities which it would need to be persuaded and to agree with the recommendation of the Reasoning part involves a high degree of anthropomorphization, and we may worry that such a theory multiplies agents or agentlike entities in undesirable ways. As we shall shortly see, there is a cluster of difficulties here concerning the theory's proliferation of agents, the explanatory value of its partitioning, and the theory's implications for the unity of the self.

In sum, the *Republic*'s theory of akratic action has serious problems and flaws and these difficulties, as we have seen, stem from the dependence of the theory of akratic action on the *Republic*'s Partition Theory. Since I think that the *Laws* abandons the theory of parts of the soul and offers a new and more attractive theory of akratic action, let me close this section by pointing to some of the more general problems facing the *Republic*'s theory of parts of the soul. Noting these general problems will allow us to see some further motivation Plato may have had for rethinking his position.

(1) As the extensive secondary literature shows, it is difficult to state the Principle of Opposites—the principle on which the partitioning of the soul depends—in such a way that Plato turns out to be making a valid and plausible argument for the conclusion he draws.<sup>24</sup> There are two especially important problems here: first, even if we grant that conflicts of desire show that the soul is divided into parts, it is not clear that Plato is entitled to the conclusion that there are only

three parts of the soul. Conflicts of desire may establish the division of the soul, but Plato does not make it clear why it is not the case that every conflict of desire establishes two new parts of the soul and certainly does not satisfactorily show that the facts about conflicts of desire produce only these three parts. A second and related problem is that Plato offers a rich characterization of each of the parts of the soul without ever providing clear justification for these characterizations. Again, even granting that conflicts of desire establish that the soul has parts, it does not establish the full description of the parts that Plato gives. We still need further argument to justify (a) the grouping of different desires together into one part, (b) the attribution to each part of some particular end or set of ends (580D-581B), and (c) the attribution to each part of some awareness of what goes on in the other parts and outside the agent.<sup>25</sup>

(2) We may be skeptical about the explanatory power of any psychological theory that divides the agent into agentlike parts. Our skepticism may range from extreme—no theory invoking homunculi can satisfactorily explain any psychological phenomenon—to the more moderate claim that homunculi of the complexity that the Partition Theory requires cannot satisfactorily explain akratic action. This objection raises deep issues that are still the subject of debate in the philosophy of mind.<sup>26</sup> But in the case at hand, we have already seen that there are concerns about the explanatory value of the parts of the soul. In the Force Model, the partitioning seemed to do no work in explaining akratic action. In some versions of the Command Model, the problems were worse: the possibility of akratic action reappeared within the parts and Plato faced the threat of a regress. In other versions, the parts either seemed to do no work or the account seemed to require *ad hoc* hypotheses.

(3) In addition to the ontological concern that the Partition Theory multiplies agents beyond necessity and the related concerns just noted about the explanatory power of the theory, partitioning presents a problem for the unity of the self. Can we give any account of the unity of the person compatible with the division of the soul into agentlike parts? How can we, compatibly with the Partition Theory, account for what seems to be the obvious fact that Leontius' desire to look at corpses and his shame and anger at the thought of doing so are both *his*?<sup>27</sup> This concern, too, raises large issues which I cannot deal with here. Plato will face the problem of accounting for both the apparent unity of a person at a time as the subject of different mental states and the apparent unity of the person over time. We cannot, of course, simply assume that since Plato's view seems to deny the strict unity of the person at a time that his view must be false. Indeed, sev-

eral recent philosophers, including Daniel Dennett and Derek Parfit, have offered strong arguments which undermine both sorts of unity and have also offered theories to explain the apparent unity of the person.<sup>28</sup> But it is not clear how much comfort Plato can draw from these theories, since, unlike Dennett and Parfit, Plato is committed to the strict unity of at least one subagent. He is committed at least in the case of the Reasoning part, and perhaps in the case of the other parts as well, to the idea that a part of the soul is a subject of experiences which is a “separately existing entity, distinct from a brain and a body, and a series of physical and mental events.”<sup>29</sup> Once we accept this claim, it will be hard to explain the apparent unity of the person who is a compound of such parts (or a compound which includes one such part). Note in this regard that Plato himself slips at times into speaking of the person as something over and above the three parts of the soul.<sup>30</sup> But the *Republic*’s Partition theory leaves no room for the person as an entity over and above the compound of the parts.

Having seen the problems of the *Republic*’s account, let us now turn to the *Laws*.<sup>31</sup>

### 3. The *Laws*’ New Account of Akratic Action

In the *Laws*, Plato accepts the possibility that an agent can both know that one course of action is better for her, ATC, and know that it is open to her to perform that action, and yet, without being physically compelled, take what she realizes is a different and worse course of action. I shall call this a case of ‘strict akratic action’.<sup>32</sup> Plato also accepts the different possibility that one’s knowledge of what is best, ATC, can be temporarily lost or obscured. I shall call this a case of “near akratic action” since the agent does not act contrary to an ATC judgment that she fully possesses at the time of the action. What we would like to know is exactly what Plato thought goes on inside an agent in cases of strict akratic action and near akratic action.<sup>33</sup>

The clearest description of strict akratic action occurs in Book 1 of the *Laws*:

Let us suppose that each of us living creatures is an ingenious puppet of the gods—whether contrived as a plaything of theirs or for some serious purpose, we do not know. But this we do know, that these affections in us [ $\tauὰ πάθη ἐν ἡμῖν$ ], like sinews or cords draw us along [ $σπῶσιν τε ἡμᾶς$ ], and being opposed to each other, pull one against the other [ $ἀνθέλκουσιν$ ] to opposite actions; and herein lies the divid-

ing line between virtue and vice. For, as our argument declares, there is one of these pulling forces [ $μιᾶ . . . τῶν ἔλξεων$ ] which every man ought always to follow [ $συνεπόμενον$ ] and in no way abandon, thereby pulling against the other sinews: it is the golden and sacred pull of calculation [ $τὴν τοῦ λογισμοῦ ἀγωγήν$ ], called the common law of the city. The other cords are hard and iron and have every sort of shape, while this one is soft since it is of gold. With that most noble pull of the law we should always cooperate; for since calculation is noble, but gentle rather than forceful, its pull needs helpers to assure that the golden kind within us may always vanquish the other kinds. (*Laws* 644D7-645B1)<sup>34</sup>

Plato tells us that he introduces this metaphor to clarify the claim that good people are able to rule themselves ( $ἄρχειν αὐτῶν$ ), but the bad are not, (644B6-7) and to clarify the meaning of the phrase “being stronger (or weaker) than oneself” ( $τὸ κρείττω ἐαυτοῦ καὶ ἥττω$ ) (645B1-3). Since failing to rule oneself and being weaker than oneself are standard descriptions of akratic action for Plato (e.g., *Rep.* 430E5-431B7), we have Plato’s own assurance that the puppet image is designed to make clear what goes on in an agent when she acts akratically and when she successfully resists.

Let us now try to fill out the details of this account. First, what sort of affections ( $πάθη$ ) does Plato have in mind here? Shortly after the above passage, Plato goes on to claim that drunkenness will affect the puppet by making its pleasures, pains, spirited emotions ( $θυμούς$ ), and loves ( $έρωτας$ ) more intense (645D6-7). On the other hand, drink will take away its perceptions, memories, opinions, and knowledge ( $φρονήσεις$ ) (645E1-2). It seems reasonable to think that the former are some of the “hard and iron” cords while the latter are among the golden cords; that is, they are instances of calculation or are intimately related to calculation. A second relevant passage claims that we can be said to be weaker or stronger with regard to pleasure and spiritedness and goes on to spell out what this means: being weaker includes the “tyranny in the soul” of spirited emotion, fear, pleasure and pain, envy and desire (863E6-8).<sup>35</sup> This gives a fairly rich account of the affections which can lead to akratic action. And in these passages spirit ( $θυμός$ ) is listed right along with various occurrent mental states, such as pleasure and pain: there is no suggestion here that spirit is agentlike.<sup>36</sup> What sort of affection is the golden cord of calculation? Calculation, according to 644D1-2, is an opinion concerned with “better and worse”; thus we can construe the agent’s calculation as an ATC judgment about what is best for the agent.

Second, what is the relation between these affections and the agent's actions? Can, for example, calculation, fear, or spirited emotion "lead" or draw a person to perform a certain action without being combined somehow with a desire? Plato does not address this issue, but some of the affections he mentions simply are desires: e.g., ἐπιθυμία and ἔρως. For other affections which are not desires, it is plausible that each instance of the affection is associated with an appropriate desire. For example, the affection of envy might be associated with a desire to improve one's own position vis-à-vis others (or diminish that of others). And with regard to calculation, we can associate with the agent's ATC judgment the desire for what she believes best for herself ATC.

Finally, these affections produce or help to produce action by "drawing" the agent to various actions. Thus here in the *Laws* Plato appears to be committed to what we would describe as a causal theory of action in which these affections are among the causes of action. The agent who can intervene in the contest also seems to have some causal role.

With these clarifications in hand, we can now see what sort of entities Plato in the *Laws* thinks we need to recognize in order to explain akratic action: all that we need to recognize are certain affections that are in the agent and the agent herself. All of the affections seem to be occurrent mental states of some sort: they are occurrent beliefs, desires, or emotions. In any case, none is described as a part and none is described as agentlike: none, for example, is described as having beliefs or desires.<sup>37</sup> The *Laws*' account of akratic action, unlike the *Republic*'s, does not invoke agentlike parts.

What then are the details of the *Laws*' account? It should run as follows. One of the iron cords, say, a desire, pulls an agent to perform a certain action X. At the same time, she knows that action Y is better ATC than any other option and that she can do either X or Y, but cannot do both X and Y. Her desire to do X pulls her more strongly than her calculation that Y is ATC best (or the desire for Y that is associated with her calculation). Thus the agent knows that Y is better ATC than X, but is more strongly motivated to do X. This case shows that we must distinguish the agent's evaluation of an option from the strength of her desire or motivation to perform that option. Her desires can vary in strength, and the strength of her desire for an option is not always directly proportional to her judgment of its goodness. And note that Plato accepts these claims while locating all the beliefs and desires involved within a single agent: none is located within an agentlike part of the soul. To continue the story: if nothing else happens, the agent will do X or try to do X. This, if it

takes place, would be a strict akratic action. But there is still one further entity that is relevant to the outcome. The agent may follow (συνέπεσθαι) the golden cord of calculation and pull along with it against the others (*Laws* 644E5-6). Less metaphorically, the agent can somehow intervene and increase the strength of her calculation or its associated desire. This, at least in some cases, will bring it about that the agent does Y. It seems reasonable to suppose that sometimes the agent's counterpull will be insufficient and she will still do X. This would again be a strict akratic action.

This account of akratic action is, I suggest, a clear improvement over that offered in the *Republic*. The problematic agentlike parts of the *Republic* are gone and we are left with an account that refers only to more familiar entities: beliefs, desires, emotions, and the agent who has them. And by distinguishing the pull or strength of a desire from the agent's evaluations of her options and relying on the notion of the strength of a desire to explain akratic conflict, Plato has made a promising beginning on what seems to be the real puzzle of akratic action: how can a *single* agent realize that one course of action is better than another and yet still knowingly do the worse?

I do want to stress, however, that this is only a beginning and that the *Laws*' theory has its own puzzles to solve. But I hope that these will seem to be more fruitful and substantive puzzles than those uncovered in our analysis of the *Republic*.<sup>38</sup> Let me mention only a few.

(1) How does the person intervene? From Plato's description, it sounds as if the person directly intervenes to increase the strength of her calculation or the strength of the desire associated with the calculation. But is such direct control of the motivational strength of one's desires really possible? If not, what indirect strategies might the agent adopt? Perhaps here we might try to identify the person's intervention with making a decision, forming an intention or exercising "the will." In this case, we shall need some account of the relations between this process and the person's ATC judgment and corresponding desire for what is best ATC.

(2) Is it always possible for the agent, if she tries hard enough, to intervene successfully? Or are at least some desires irresistible? How many failures of self-control are due to simply irresistible desires? And if, due to a bad upbringing and excessive temptation, most agents are unable to resist, how can they be blamed for their failure to exercise self-control?<sup>39</sup>

(3) What determines whether the agent even tries to intervene, and is it in some way necessary that she intervene on the side of her calculation or desire for what is ATC best? Does she usually or ever undertake some kind of deliberation to try to decide whether or not to

intervene? If so, what is such deliberation like? And if the agent's deliberation ends with the formation of a belief that it is ATC best for her to intervene, is it possible that she might akratically fail even to try to intervene?

(4) Important problems still surround the relation between the strength of a desire and the content of the agent's desires and beliefs. For example, it seems to be quite common for the recognition that one course of action is ATC best to lead to the diminishment of the strength of desires for incompatible options and to an increase in the strength of my desire for what I now realize to be ATC best. But how is this possible? Can we give any further explanation of the fact that the strength of a person's desire is sometimes, but only sometimes, sensitive to the desire's own content and the content of the agent's other beliefs?<sup>40</sup>

Plato does not face these questions in the *Laws* and I cannot try here to tease out the answers he might give, but before turning to the next section, let me briefly discuss his account of near akratic action. Immediately after our passage, Plato sketches a slightly different possibility (*Laws* 645D1-646A5). Here the story runs as follows. Drink increases the "intensity" of certain desires and pleasures. (It makes them *σφοδροτέρας* [645D6-7].) At the same time, drink causes a person's opinions, memories, and so on to abandon her (*ἀπολεῖπεται*). This sort of account could explain a case of apparent strict akratic action in the following way. Again suppose that I know that Y is ATC the best option for me and that doing Y is incompatible with doing X. Suppose I also have a desire to do X. Drink may increase the strength of my desire to do X and cause me to lose my ATC belief. I then do X without an ATC belief. (Why does drink not also increase my desire to do Y? Plato could argue that this desire is dependent on a belief arrived at by reasoning. As I drink, I lose my grip on this reasoning and thus lose the belief that would issue from it. Therefore, I also lose the desire dependent on this belief.)

This is not a case of strict akratic action since the agent does not act contrary to an ATC belief that she possesses at the time of the action. But it is near to akratic action in that the agent, without coming to believe that another course of action is ATC best, acts contrary to the last ATC judgment that she had. We might wonder what role the person has in this account: she might try to intervene to preserve the ATC judgment, but by the time it is lost, there is no conflict left in which the person can intervene.<sup>41</sup> The basic idea behind this account of akratic action is that apparent instances of strict akratic action are analyzed in such a way that the agent does not act

against an ATC judgment that she fully possesses at the time of the action. This is a strategy that Plato himself has employed before in the *Republic*, and it is later employed by Aristotle in his own analysis of akratic action.<sup>42</sup>

#### 4. Parts of the Soul in the *Laws*

I have argued that the *Laws*' explanation of akratic action does not invoke parts of the soul, and I have made the more general claim that the Partition Theory is absent from the moral psychology of the *Laws*. Let me now briefly provide some support for the latter claim. First, the logical basis of partitioning and the terminology of partitioning are absent from the *Laws*. The Principle of Opposites which is essential to the *Republic*'s argument for the partitioning of the soul is never stated or referred to in the *Laws*; and neither *τὸ λογιστικόν*, nor *τὸ θυμοειδές*, nor *τὸ ἐπιθυμητικόν* is found in the *Laws*. (Below I shall consider Plato's terms for "part.") Of course, the mere absence of these two features is not decisive, since they are more or less absent from the *Phaedrus* and the *Timaeus*, two dialogues in which Plato is often thought to have divided the soul into agentlike parts.<sup>43</sup> But in the *Timaeus* and perhaps the *Phaedrus*, Plato uses other means to make it clear that he is dividing the soul and attributing some sort of agency to the parts.

In the *Phaedrus*, we find a long and elaborately worked out metaphor in which the soul of a human being is conceived of as a charioteer and two winged horses. Although it is difficult to fix precisely the literal meaning of this metaphor, it at least suggests the essentials of the Partition Theory: the soul is seen as a compound of three entities, and each of these entities is itself an agent of some sort. But the metaphor should not be pressed too hard and cannot be taken as firm evidence that Plato still held the Partition Theory at the time of the *Phaedrus*. And it may be significant that the partitioning is not presented as a plain piece of theorizing, however tentative, but as a story or myth (*μῦθος*).<sup>44</sup> In any case, this myth or image does not occur in the *Laws*.

The *Timaeus* distinguishes three classes or kinds of thing (*εἴδη* or *γένη*) within the individual soul.<sup>45</sup> In particular, an immortal kind or principle (*ἀρχή*) within the soul is distinguished from a mortal kind (*Tim.* 69C3-70A2). The mortal kind itself is divided into two parts (69E3 ff.): a better and a worse part which roughly correspond to the Spirited part and the Desiring part (cf. 70D7). And various psycho-

logical attitudes and activities are freely attributed to these parts (69E3-73A8 *passim*). The *Laws*, on the other hand, has no mention of such εὐδη or γένη within a single soul. The closest it comes are the phrase at 691C6 “the nature of a mortal soul” (θνητῆς ψυχῆς φύσις), which is simply an elevated expression for the soul of a mortal, and references at 713C6 to “human nature” (ἡ ἀνθρωπεία φύσις) and to “mortal nature” (ἡ θνητὴ φύσις) at 875B7, but these are nothing more than ways of referring to the psychological character that humans, as opposed to gods, have and do not commit Plato to the idea that the soul is composed of agentlike parts.<sup>46</sup>

There is one further noteworthy feature of the *Timaeus* account. We have seen that in the *Republic* Plato is committed to the idea that one part of the soul can persuade another part and, in particular, to the idea that the Reasoning part can persuade the Desiring part. This idea occurs several times in the *Republic*, and in Book 8 Plato uses this distinction to help to distinguish the philosopher from the oligarchic man (the money-lover). The oligarchic man, Plato says, can control some of his desires in order to satisfy his long-run goals—e.g., he can refrain from having a bottle of gin before going off to the Stock Exchange. But the oligarchic man is defective because he controls his Desiring part by force, not by reason. In the oligarchic man’s public dealings

where he enjoys the reputation of being just, his better part keeps down his other evil desires by force [βίᾳ]; he does not persuade them that it is not better [ὅτι οὐκ ἀμεινον] [sc. that they be satisfied], nor does he make them gentle by reasoning [λόγῳ] with them, but he suppresses them by compulsion and fear [ἀνάγκῃ καὶ φόβῳ] (*Rep.* 554C11-D3).<sup>47</sup>

For this criticism of the oligarchic man to have a point, Plato must think that the philosopher can persuade his Desiring part by argument, by means of λόγοι, that it is better for it to go along with reason. Thus the sort of judgment that the Desiring part ends up making here is an ATC judgment: the Desiring part initially wants to satisfy certain desires but is persuaded by the Reasoning part that, in light of other considerations, it is better not to do so. Since the range of information open to the Desiring part is limited (it cannot judge of what is best ATC for the whole soul or have knowledge of the Forms), it cannot take into account all possible considerations. But if the Reasoning part’s persuasion is to work, the Desiring part must take more than one consideration into account; and we can describe this as an

ATC judgment in which the Desiring part takes into consideration what it can. Weighing different considerations, even if it is done briefly and not all are taken into account, is sufficient for the essential idea of an ATC judgment.<sup>48</sup>

But we see in the *Timaeus* that Plato has second thoughts about this attribution of reasoning capacities to the Desiring part. And in the *Timaeus*, Plato modifies his theory: the Desiring part is now, Plato thinks, unable to understand or respond to reason, but can respond to images and appearances (εὐδώλα and φαντάσματα) generated by the Reasoning part.<sup>49</sup> It is especially striking that Plato explicitly rejects the idea of the *Republic* passage: the Reasoning part here can affect the Desiring part only by frightening it (φόβοι 71B5) or making it feel cheerful and serene (ἴλεών τε καὶ εὐήμερον 71D1-2). Plato does not tell us what sort of cognitive capacities are necessary to understand and respond to images, but what is important to note is his growing reluctance to treat the Desiring part as fully agentlike.

Next, let us consider Plato’s word for “part”: μέρος.<sup>50</sup> The only relevant occurrence of the term in the *Laws* comes while Plato is trying to reconcile the commonsense legal distinction between willing (έκών) and unwilling (άκων) wrongdoing with his own view that all wrongdoing is unwilling. Here Plato refers to spirit (θυμός) as “either an affection [πάθος] or a part (μέρος) of the soul” (863B2-3) and does not try to choose between the two options: doing so is not necessary for settling the issue at hand in this passage. But, as we have seen, there is one place in the *Laws* where Plato does try to clarify what is involved in akratic action, that is, in the puppet passage (644D7-645B1). And in that passage, he is careful to refer only to the agent and the agent’s affections (πάθη); he never refers to parts (μέρη) of the soul.

We find two very brief references in the *Laws* to what might at first be thought to be agentlike parts of the soul. The first is the phrase τό τῆς τοιαύτης ἀρχῆς γιγνόμενον ὑπήκοον ἐκάστων at 864A5-6. Even if this is a reference to some entity or entities in the soul—“whatever thing in each individual which becomes obedient to such a principle”—we do not learn anything about these entities (How many are there? Do they have beliefs and desires or are they beliefs and desires?) and cannot take this to commit Plato to agentlike parts of the soul. But also note that Plato has just said that the distinction between a part (μέρος) and an affection (πάθος) of the soul is not relevant to his present discussion (863B1-5) and has emphasized be-

fore and after our phrase that he is not concerned with terminological exactitude (863E5-6 and 864A8-B1). Thus we should not press his phrasing too hard. Finally, we need not translate  $\tau\ddot{o}$  . . .  $\bar{\nu}\pi\bar{\eta}\chi\bar{o}\bar{o}\bar{o}$  as “the thing” or “the part.” Instead we can take it as an abstract substantive and translate “everything which takes place in obedience to such a principle in every individual” or “all obedience of individuals to such a principle.”<sup>51</sup> The second passage is a reference at *Laws* 689A9 to  $\tau\ddot{o}$   $\pi\lambda\bar{\eta}\bar{\theta}\bar{o}\bar{s}$   $\tau\bar{\eta}\bar{s}$   $\psi\bar{\chi}\bar{\eta}\bar{s}$  (the “bulk” or “majority of the soul”) which Plato identifies with  $\tau\ddot{o}$   $\lambda\bar{\nu}\pi\bar{o}\bar{u}\bar{m}\bar{e}\bar{n}\bar{o}\bar{v}$   $\kappa\bar{\alpha}\bar{l}\bar{i}$   $\bar{\eta}\bar{\delta}\bar{o}\bar{m}\bar{e}\bar{n}\bar{o}\bar{v}$  (“that which feels pain and pleasure” 689B1). Once again, we do not learn anything about  $\tau\ddot{o}$   $\pi\lambda\bar{\eta}\bar{\theta}\bar{o}\bar{s}$  (or whatever minorities are left over): for example, we are not told how many things are involved or whether any of them have beliefs or desires. (Cf. *Rep.* 580D11-581A1 where “the greatest thing [μέγιστον]” in the Desiring part is its desires for food and so on taken collectively.) Nor can we identify  $\tau\ddot{o}$   $\pi\lambda\bar{\eta}\bar{\theta}\bar{o}\bar{s}$  here with the Desiring part of the *Republic* (cf. *Rep.* 442A5-6 and 431A3-B2), since here  $\tau\ddot{o}$   $\pi\lambda\bar{\eta}\bar{\theta}\bar{o}\bar{s}$  is identified with  $\tau\ddot{o}$   $\lambda\bar{\nu}\pi\bar{o}\bar{u}\bar{m}\bar{e}\bar{n}\bar{o}\bar{v}$   $\kappa\bar{\alpha}\bar{l}\bar{i}$   $\bar{\eta}\bar{\delta}\bar{o}\bar{m}\bar{e}\bar{n}\bar{o}\bar{v}$ , but in the *Republic* all three parts have their own pleasures. Finally, Plato does not explain the “ignorance” that he is discussing at *Laws* 689A1-E2 as a cognitive failure on the part of  $\tau\ddot{o}$   $\pi\lambda\bar{\eta}\bar{\theta}\bar{o}\bar{s}$ , rather this ignorance simply consists in the fact that the person feels pleasure in things which she knows are not best for her (cf. *Laws* 696C8-10). And this is the last that we hear of  $\tau\ddot{o}$   $\pi\lambda\bar{\eta}\bar{\theta}\bar{o}\bar{s}$  of the soul in the *Laws*.

But, finally, perhaps the most important consideration is that the parts of the soul do not do any philosophical work in the *Laws*. Akrasia is explained and a strategy for avoiding it outlined which does not commit Plato to the existence of the parts of the soul. Plato defines the virtues and argues for the claim that the most virtuous life is the most pleasant life again without invoking the parts of the soul. Yet even if I am wrong and Plato did continue to believe in parts of the soul, even if they are merely offstage in the *Laws*, my fundamental point still holds firm: they are no longer thought by Plato to be necessary for framing or solving the philosophical problems surrounding akratic action.

Although we cannot determine with certainty what led Plato to abandon the theory of parts of the soul, we have seen that there are a number of serious problems for this theory and the theory of akratic action which invokes the parts of the soul. The abandonment of the Partition Theory would be a reasonable response to these difficulties,

and in the *Laws* we see Plato formulating a theory of akratic action which is not committed to agentlike parts of the soul.<sup>52</sup>

## 5. The Reciprocity of the Virtues and Political Pessimism

We have seen that Plato's theory of akratic action has many connections with important issues in his philosophy of mind. But it is also tightly connected with significant issues in Plato's moral and political philosophy, and I shall conclude by considering two of them. First, Plato's acceptance of the possibility of akratic action presents a threat to one important tenet of his early moral theory, and in the *Laws* we see Plato trying out one possible response. The thesis in question is the Reciprocity of the Virtues: A person has a virtue if and only if she has every virtue.<sup>53</sup> Consider the following argument against this thesis: (1) Some people know that acting virtuously is ATC best for them and still act akratically. (2) If a person acts akratically, then she lacks at least one virtue. (3) If a person knows that acting virtuously is ATC best for her, then she has the virtue of wisdom. From (1) & (3): (4) Some people have the virtue of wisdom and still act akratically. From (4) & (2): (5) Some people have the virtue of wisdom and still lack at least one virtue. (5) is inconsistent with the Reciprocity of the Virtues.

Plato in the *Laws* accepts (1), for example at 689A1-C3, 874E8-875C3 and 902A6-B2. With regard to (2), whether we construe moderation as the ability to act in accordance with one's ATC judgment of what is best, despite some contrary desires, or as the state in which all of one's desires are for what one believes to be ATC best, (2) will still be true since the person who acts akratically lacks at least the virtue of moderation.<sup>54</sup> Nor can we avoid (5) by allowing that a person might occasionally act akratically and still possess all the virtues: we can simply modify (1) and (2) by the appropriate insertion of 'often'.<sup>55</sup> We might, finally, try to reject (3). One way of doing so is by claiming that knowing that acting virtuously is ATC best for you is not sufficient for possessing the virtue of wisdom: one has to know other things. But even if we accept this suggestion, it is not enough to remove the threat to the Reciprocity thesis. Merely increasing the amount of knowledge required for wisdom does not guarantee that those who meet the new higher requirement do not act akratically. Indeed, the passages supporting (1) and the analysis of akratic action we have just gone through seem to show that no amount of knowledge

is sufficient to guarantee that the agent will act in accordance with her ATC judgment: her desire for something else may just be too strong.

So must we accept (5) and the denial of the Reciprocity of the Virtues?<sup>56</sup> No, since there is another way to avoid (3) and thus (5), and it is one that Plato explores in the *Laws*.<sup>57</sup> The basic idea is that something more than knowledge is needed in order to have the virtue of wisdom, and Plato at least flirts with this idea at 689A1 ff. Unfortunately, the passage is neither well worked out nor well integrated with the rest of the *Laws*. Here Plato claims that “the ultimate and greatest stupidity [ἀμαθίαν]” is “when someone does not love, but rather hates, what in his opinion is noble or good, and loves and welcomes what in his opinion is wicked and unjust.” Plato goes on to describe this condition as a “dissonance [διαφωνίαν] between pleasure and pain, and the opinion that is according to reason,” and extends the point to include knowledge: “when the soul opposes knowledge, or opinions, or reason [δταν ούν ἐπιστήμαις ή λόγῳ ἐάτιωται]—the natural rulers—this I call lack of intelligence” (689A5-B4). On the other hand, “the finest and greatest of consonances [συμφωνιῶν] is the greatest wisdom [σοφία]” (689D6-7).<sup>58</sup>

This passage at 689A1 ff. suggests that Plato is aware of the threat that akrasia poses to the Reciprocity of the Virtues and it sketches one strategy for dealing with this threat, which is to make the following two conditions individually necessary and jointly sufficient for wisdom: (a) knowledge of what is ATC best for oneself,<sup>59</sup> and (b) consonance. Since a person who possesses (a) may still act akratically, the idea is to construe (b) in such a way that a person who possesses both (a) and (b) will not act akratically. Since Plato does not make the concept of consonance entirely determinate, we must go a little beyond the text. The weakest satisfactory definition of consonance would be that a person who possesses consonance always most strongly desires to do what she believes is ATC best. Such consonance, it seems, may or may not require the sort of intervention by the agent described above.

Some of the other details of Plato’s proposal remain obscure, but we now have enough of the idea to evaluate it. This broadening of the concept of wisdom is not a policy that Plato himself sticks to in the *Laws* (e.g., 863A7 ff. and 631C5 ff.) and it seems to be merely a formal and unsatisfying resolution of the problem that Plato is facing. The Reciprocity of the Virtues, on first inspection, seems to be a substantive and controversial thesis, and the acceptance of akrasia seems to present a serious challenge to it. Resolving the threat to the Reciprocity of the Virtues by broadening the concept of wisdom in

the way just suggested does remove the threat but only by making (3) trivially false and thus blocking the inference to (4) and (5). Indeed, as long as consonance is sufficient for moderation (whether understood as complete harmony or self-control), having wisdom trivially entails possessing every other virtue (cf. 631C5 ff. and 633C8-D3). Although Plato sees more clearly in the *Laws* than elsewhere that there may be a conflict between the acceptance of akrasia and the Reciprocity of the Virtues, his proposed solution merely robs this thesis of much of its interest.

This problem for the Reciprocity of the Virtues stems simply from Plato’s acceptance of the possibility of akratic action, not from the new analysis of it he provides in the *Laws*. I would like to close by briefly noting the connection between his new analysis of akratic action and one significant development in his political philosophy.<sup>60</sup> Plato is more pessimistic in the *Laws* than in the *Republic* about the capacity of those who hold absolute power to remain uncorrupted: in the *Laws*, Plato thinks that absolute power will tend to make its possessors act akratically, even if they start off as just people, and even if they have been philosophically trained (875A1 ff., cf. 691C1 ff. and 713C2 ff.).<sup>61</sup> This is one reason for the increased concern in the *Laws* for safeguards in the distribution of political power. What can explain Plato’s increased pessimism?

Gregory Vlastos suggests that Plato’s opinion was changed by his “final encounter with Dionysus the Younger, when he saw the ugly face of autocratic power at closer and more painful quarters than at any other time in his life.”<sup>62</sup> But this explanation is unsatisfying. Even if Plato had thought that Dionysus was both well brought up and had attained full philosophical understanding of virtue and happiness, it would be rash to generalize from this one case to all possible cases. But Plato thought that Dionysus was badly brought up (*Epist. VII* 332CD) and had a low opinion of his philosophical knowledge (*Epist. VII* 338D-339B, 340B-341B, 344D-345C). It would thus be irrational for Plato to generalize from his experience with Dionysus to the capacities of all human beings. Dion would come closer to Plato’s ideal (although his early upbringing must have been somewhat defective, *Epist. VII* 327B1-4) and Vlastos rightly notes that “Plato did not look on Dion as an autocrat corrupted by power but as a good man who was undone because he underestimated the villainy of his associates.”<sup>63</sup>

I suggest that we look instead to the shift I have sketched in Plato’s views on akrasia and moral psychology.<sup>64</sup> In the *Republic*, as we have seen, Plato believed that the two lower parts of the soul could be brought to accept the rule of reason and to be obedient to it. Despite the serious difficulties we have seen surrounding this idea, Plato

may think that once the lower two parts are on reason's side, they will no longer prompt the agent to act akratically. Thus a philosopher king will continue to act justly even when he possesses absolute, unchecked power. In the *Laws*, however, there are no agentlike lower parts of the soul to be persuaded, and the only ultimate protection against akratic action is the individual's power of self-control. And although Plato does not answer all the questions he needs to answer, he clearly thinks that in the circumstance of having absolute power, the vast majority of individuals will have powerful but unjust desires and will fail or be unable to intervene in the appropriate way. Once the possibility of internal persuasion is gone, the only safeguard left is the more precarious one of self-control.<sup>65</sup>

## Bibliography

Adam, J. 1979. *The Republic of Plato*. Cambridge: Cambridge University Press.

Annas, J. 1981. *An Introduction to Plato's Republic*. Oxford: Clarendon.

—. 1982. "Plato's Myths of Judgment." *Phronesis* 27: 119-43.

Belfiore, E. 1986. "Wine and *Catharsis* of Emotions in Plato's *Laws*." *Classical Quarterly* 36: 421-37.

Bloom, A. 1968. *The Republic of Plato*. New York: Basic Books.

Brandwood, L. 1976. *A Word Index to Plato*. Leeds: W. S. Maney & Son.

—. 1990. *The Chronology of Plato's Dialogues*. Cambridge: Cambridge University Press.

Bratman, M. 1987. *Intention, Plans, and Practical Reason*. Cambridge, Mass.: Harvard University Press.

Burge, E. 1971. "The Ideas as *Aitiai* in the *Phaedo*." *Phronesis* 16: 1-13.

Burnyeat, M. 1976. "Plato on the Grammar of Perceiving." *Classical Quarterly* 26: 29-51.

—. ed. 1990. *The Theaetetus of Plato*. Indianapolis, Ind.: Hackett.

Bury, R. 1926. *Plato: Laws*. New York: G. P. Putnam's Sons.

—. 1975. *Plato: Timaeus, Critias, Cleitophon, Menexenus, Epistles*. Cambridge, Mass.: Harvard University Press.

Charlton, W. 1988. *Weakness of Will*. Oxford: Basil Blackwell.

Crombie, I. 1969. *An Examination of Plato's Doctrines*. Vol. 1. London: Routledge & Kegan Paul.

Davidson, D. 1980. *Essays on Actions and Events*. Oxford: Oxford University Press.

—. 1980a. "Actions, Reasons, Causes." In Davidson 1980, 3-19.

—. 1980b. "How Is Weakness of the Will Possible?" In Davidson 1980, 21-42.

—. 1982. "Paradoxes of Irrationality." In R. Wollheim and J. Hopkins, eds. *Philosophical Essays on Freud*. Cambridge: Cambridge University Press, 289-305.

—. 1986. "Deception and Division." In J. Elster, ed. *The Multiple Self*. Cambridge: Cambridge University Press, 79-92.

Dennett, D. 1981. "Intentional Systems." In D. Dennett, *Brainstorms*. Cambridge, Mass.: MIT Press, 3-22.

—. 1991. *Consciousness Explained*. Boston, Mass.: Little, Brown.

England, E. B. ed. 1921. *The Laws of Plato*. 2 vols. Manchester: Manchester University Press.

Ferejohn, M. 1984. "Socratic Thought-Experiments and the Unity of Virtue Paradox." *Phronesis* 29: 105-22.

Ferrari, G. 1985. "The Struggle in the Soul: Plato, *Phaedrus* 253C7-255A1." *Ancient Philosophy* 5: 1-10.

—. 1987. *Listening to the Cicadas*. Cambridge: Cambridge University Press.

Fine, G. 1988. "Owen's Progress: A Critical Notice of *Logic, Science and Dialectic*." *Philosophical Review* 97: 373-99.

Fodor, J. 1981. "The Appeal to Tacit Knowledge in Psychological Explanations." In J. Fodor, *Representations*. Cambridge, Mass.: MIT Press, 63-78.

Gauthier, D. 1968. "The Unity of Wisdom and Temperance." *Journal of the History of Philosophy* 6: 157-59.

—. 1971. "Comments on Hare." In R. Binkley et al., eds. *Agent, Action, and Reason*. Toronto: University of Toronto Press, 98-108.

Gill, C. 1985. "Plato and the Education of Desire." *Archiv für Geschichte der Philosophie* 67: 1-26.

Gosling, J. 1990. *Weakness of the Will*. London: Routledge & Kegan Paul.

Guthrie, W. 1987. *A History of Greek Philosophy*. Vol. 4. Cambridge: Cambridge University Press.

Hackforth, R. 1946. "Moral Evil and Ignorance in Plato's Ethics." *Classical Quarterly* 40: 118-20.

Hare, R. 1971. "Wanting: Some Pitfalls." In R. Binkley et al., eds. *Agent, Action, and Reason*. Toronto: University of Toronto Press, 81-97.

Heath, M. 1989a. "The Unity of Plato's *Phaedrus*." *Oxford Studies in Ancient Philosophy* 7, 151-73.

—. 1989b. "The Unity of the *Phaedrus*: A Postscript." *Oxford Studies in Ancient Philosophy* 7, 189-91.

Irwin, T. 1979. *Plato's Moral Theory*. Oxford: Clarendon.

Joseph, H. 1971. *Essays in Ancient and Modern Philosophy*. Freeport, N.Y.: Books for Libraries Press.

Korsgaard, C. 1989. "Personal Identity and the Unity of Agency: A Kantian Response to Parfit." *Philosophy & Public Affairs* 18:101-32.

Ledger, G. 1989. *Re-counting Plato*. Oxford: Clarendon.

Liddell, H., Scott, R. and Jones, H. 1978. *A Greek-English Lexicon*. 9th ed. Oxford: Clarendon.

Mele, A. 1987. *Irrationality*. Oxford: Oxford University Press.

—. 1990. "Errant Self-Control and the Self-Controlled Person." *Pacific Philosophical Quarterly* 26: 19-30.

—. 1992. *Springs of Action*. Oxford: Oxford University Press.

Moline, J. 1978. "Plato on the Complexity of the Psyche." *Archiv für Geschichte der Philosophie* 60: 1-26.

—. 1981. *Plato's Theory of Understanding*. Madison: University of Wisconsin Press.

—. 1988. "Plato on Persuasion and Credibility." *Philosophy and Rhetoric* 21: 260-78.

Nussbaum, M. 1980. "Shame, Separateness, and Political Unity: Aristotle's Criticism of Plato." In A. Rorty, ed. *Essays on Aristotle's Ethics*. Berkeley: University of California Press, 395-435.

—. 1986. "'This story isn't true': madness, reason, and recantation in the *Phaedrus*." In M. Nussbaum. *The Fragility of Goodness*. Cambridge: Cambridge University Press, 200-33.

O'Brien, M. 1957. "Plato and the 'Good Conscience,' *Laws* 863E5-864B7." *Transactions of the American Philological Association* 88: 81-87.

—. 1967. *The Socratic Paradoxes and the Greek Mind*. Chapel Hill: University of North Carolina Press.

Pangle, T. 1980. *The Laws of Plato*. New York: Basic Books.

Parfit, D. 1984. *Reasons and Persons*. Oxford: Oxford University Press.

Pears, D. 1971. "Comments on Hare." In R. Binkley et al., eds. *Agent, Action, and Reason*. Toronto: University of Toronto Press, 108-27.

—. 1984. *Motivated Irrationality*. Oxford: Oxford University Press.

Penner, T. 1973. "The Unity of Virtue." *Philosophical Review* 82: 35-68.

—. 1978. "Thought and Desire in Plato." In G. Vlastos, ed. *Plato II: Ethics, Politics, and Philosophy of Art and Religion*. Notre Dame: University of Notre Dame Press, 96-118.

Price, A. 1989. *Love and Friendship in Plato and Aristotle*. Oxford: Clarendon.

—. 1990. "Plato and Freud." In C. Gill, ed. *The Person and the Human Mind*. Oxford: Clarendon, 247-70.

Rees, D. 1957. "Bipartition of the Soul in the Early Academy." *Journal of Hellenic Studies* 77: 112-18.

Reeve, C. 1988. *Philosopher-Kings*. Princeton, N.J.: Princeton University Press.

Robinson, R. 1971. "Plato's Separation of Reason from Desire." *Phronesis* 16: 38-48.

Robinson, T. 1970. *Plato's Psychology*. Toronto: University of Toronto Press.

Rowe, C. 1986. "The Argument and Structure of Plato's *Phaedrus*." *Proceedings of the Cambridge Philological Society* n.s. 32: 106-25.

—. 1989. "The Unity of the *Phaedrus*: A Reply to Malcolm Heath." *Oxford Studies in Ancient Philosophy* 7, 175-88.

Santas, G. 1966. "Plato's *Protagoras* and Explanations of Weakness." *Philosophical Review* 75: 3-33.

Saunders, T. 1962. "The Structure of the Soul and the State in Plato's *Laws*." *Eranos* 60: 37-55.

Shorey, P. 1978. *Plato: Republic*. Cambridge, Mass.: Harvard University Press.

Sorabji, R. 1990. "Perceptual Content in the Stoics." *Phronesis* 35: 307-14.

Stalley, R. 1975. "Plato's Argument for the Division of the Reasoning and Appetitive Elements in the Soul." *Phronesis* 20: 110-28.

Taylor, A. 1960. *Plato: the Laws*. London: J. M. Dent and Sons.

Taylor, C. 1982. *Plato: Protagoras*. Oxford: Clarendon.

Thalberg, I. 1985. "Questions about Motivational Strength." In E. Lepore and B. McLaughlin, eds. *Actions and Events*. Oxford: Basil Blackwell, 88-103.

Vander Waerdt, P. 1985a. "The Peripatetic Interpretation of Plato's Tripartite Psychology." *Greek, Roman and Byzantine Studies* 26: 283-302.

—. 1985b. "Peripatetic Soul-Division, Posidonius, and Middle Platonic Moral Psychology." *Greek, Roman and Byzantine Studies* 26: 373-94.

Velleman, J. 1989. *Practical Reflection*. Princeton: Princeton University Press.

Vlastos, G. 1977. "The Theory of Social Justice in the *Polis* in Plato's Republic." In H. North, ed. *Interpretations of Plato*. Leiden: *Mnemosyne Suppl.*, 1-40.

—. 1981. *Platonic Studies*. 2d ed. Princeton: Princeton University Press.

—. 1981a. "Socratic Knowledge, Platonic 'Pessimism'." In Vlastos 1981 204-17.

—. 1981b. "Reasons and Causes in the *Phaedo*." In Vlastos 1981, 76-110.

—. 1981c. "The Unity of the Virtues in the *Protagoras*." In Vlastos 1981, 221-65.

—. 1981d. "Socrates on 'the Parts of Virtue'." In Vlastos 1981, 418-23.

—. 1983. "The Socratic Elenchus." *Oxford Studies in Ancient Philosophy* 1, 27-58.

Walker, A. 1989. "The Problem of Weakness of Will." *Noûs* 23: 653-76.

Walsh, J. 1963. *Aristotle's Conception of Moral Weakness*. New York: Columbia University Press.

Watson, G. 1977. "Skepticism about Weakness of the Will." *Philosophical Review* 86: 316-39.

Williams, B. 1973. "Ethical Consistency." In B. Williams, ed. *Problems of the Self*. Cambridge: Cambridge University Press, 166-86.

Woodruff, P. 1976. "Socrates on the Parts of Virtue." *Canadian Journal of Philosophy Suppl.* vol. 2: 101-16.

Woods, M. 1987. "Plato's Division of the Soul." *Proceedings of the British Academy* 73: 23-48.

## Notes

1. *Faute de mieux*, I use the standard translations, although they may be misleading: for example, all three parts, not merely the Desiring part, have desires.

2. This formulation is indebted to Annas (1981, 142) and a suggestion by one anonymous referee.

3. For desires and pleasures, see, e.g., 580D3-587E4; for  $\betaούλεσθαι$  and  $\epsilon\thetaέλειν$ , e.g., 437B1-C10 and 439A1-D2. For beliefs, e.g., 442B5-D1, 574D1-575A7; for persuasion and agreement, e.g., 442B5-D1, 554C11-E5 and 589A6-B6. Also, cf. Annas (1981, 109-52), Burnyeat (1976), Joseph (1971, 41-82 and 156-78), Moline (1978) and (1981, 52-78), and Price (1989, 1990). All three parts are capable of at least some form of means-end reasoning: 580E2-581A7, cf. Annas (1981, 129-30, 139) and Moline (1981, 59-61). The Reasoning part is also capable of making all things considered (ATC) judgments about what is best for the whole soul (*Rep.* 442C5-8); I argue below that the Desiring part, although it cannot judge of what is ATC best for the whole soul, can make true ATC judgments about what is best for it. As we shall see in section 4, in the *Timaeus* Plato comes to find the attribution of reasoning to the Desiring part problematic.

4. Annas (1981), Burnyeat (1976), and Moline (1978, 1981, 1988). Irwin (1979, 200 and 245) and Ferrari (1985, 1987) also seem to think that the parts of the soul are anthropomorphic. Cf. Price (1989, 1990).

5. At least embodied persons are complex: *Republic* 608C-612A seems to suggest that the disembodied soul is simple and consists solely of the Reasoning part, but cf. T. Robinson (1970, 50-58).

6. The idea of drawing some distinction here was suggested by Bernard Williams in the course of a 1989 seminar at Berkeley. Although I am indebted to Williams and some of the other participants for their suggestions, none should be assumed to endorse my present claims.

7. I adopt Woods' (1987, 30) name for the principle.

8. *Rep.* 439B5-6 shows that these pairs of opposites are actions ( $\piοτ\cdot\etaματα$ ). I draw on Bloom (1968) and Shorey (1978) in quotations from the *Republic* and Bury (1926) and Pangle (1980) for the *Laws*.

9. The only explicit comment on this point in the literature is by J. Adam: "One part of the soul asks, and the other answers" (1979, 1: 249). But this seems wrong. The opposites described here are done by different parts of the soul. The assent and dissent are activities of different parts and the assent and dissent each go on entirely within one part. Plato here does not explicitly argue that assent and dissent are genuine opposites, and he does not analyze the nature of their opposition. Nevertheless, he seems aware that the question concerns the compatibility of two psychological or propositional attitudes (accepting a proposition or asserting a positive imperative and rejecting a proposition or asserting a negative imperative), and he does not try to assimilate the impossibility of the

same thing doing both of these to the impossibility of the same proposition being true and false. Cf. Crombie (1969, 355-56) and Stalley (1975). On related issues, cf. Gauthier (1971), Hare (1971), and Pears (1971).

10. Cf. Mele (1987, 1990, 1992) and fn. 14. For refinements to the statement of the situation and the generalization described in the above paragraph which, e.g., take temporal indices into account, cf. Mele (1992, 46-78).

11. Thinking of desires as tugs or pulls does not require or even suggest that the desire has no intentional content. One might think that the assent or dissent specifies the content of a desire while the description of the desire as a pull merely attributes strength to the desire. But this does not seem to be Plato's idea here: even when described as a tug or a pull, the desire is directed toward some object and this may specify its intentional content (e.g., 439A9-B1). This also leaves open the possibility that a desire's content affects its strength. Cf. Velleman (1989, 190-96) on the distinction between the force a desire has independently of its significance being grasped and force it has because its significance is grasped.

12. Although this is the natural way of filling out the Force Model and seems to be the picture suggested by 439A9 ff., Plato never explicitly makes this claim. Cf. Davidson's well-known principle P1: "If an agent wants to do X more than he wants to do Y and he believes himself free to do either X or Y, then he will intentionally do X if he does either X or Y intentionally" (1980b, 23). On intentionality, cf. fn. 33.

13. Davidson (1982, 1986) and Pears (1984) offer partitioning theories which are more sophisticated than Plato's, but also less ambitious. Davidson (1982, 300, and 300 n. 6) admits that his account of the partitioning "leaves much unexplained, for we want to know why this double structure developed, how it accounts for the action taken, and also, no doubt, its psychic consequences and cure" and that "I have nothing to say about the number or nature of the divisions of the mind." Pears' broadly Davidsonian theory is intended as an account of motivated irrational belief, not akratic action. Cf. Mele (1987, 75-84 and 139-49).

14. On the difficulties surrounding the notion of the strength of a desire, see Charlton (1988, 126-33), Gosling (1990, 174-85), Mele (1987, 1990, 1992), Penner (1978), Santas (1966) and Thalberg (1985). 588E3-589A4 may suggest that the strength of the various parts of the soul affects the strength of their desires. But it is difficult to see what precise sense we can give to the strength of a soul part as opposed to the strength of a desire.

15. The relation between Plato's views and causal theories of mind has been oddly ignored in recent discussions of Plato's philosophy of mind. On Plato's notion of causation, see, e.g., Burge (1971) and Vlastos (1981b).

16. I am not claiming that explanation by reasons precludes or renders unnecessary explanation by causation, but this idea is one that many people have at times found persuasive. For a powerful attack on it which has convinced many contemporary philosophers, see Davidson (1980a).

17. Davidson (1980b, 35) may have in mind *Rep.* 553B8-D4; but we do not find the third actor in Book 4 and its presence is inconsistent with Plato's partitioning. In fairness to Davidson, interpreting Plato is not his concern in the passage.

18. If the imperatives were needed to specify the intentional content of the desire, they would have an indirect role insofar as the content of a desire affects its strength.

19. For such persuasion, see 442B5-D1, 554C11-E5, and 589A6-B6. In a moderate and just person, the Reasoning part is to persuade both of the other parts, but here I concentrate on the interaction of the Reasoning part and the Desiring part. The idea that the two lower parts can be persuaded by the Reasoning part is not easily eliminated from the *Republic*, since it seems essential to Plato's account of the virtues of moderation and justice and to his distinction between the philosopher and the oligarchic man (554C-E).

20. If, when it is not persuaded by the Reasoning part, the Desiring part has an ATC judgment that it is best to act in accordance with its own strongest desire and this judgment conflicts with the ATC judgment of the Reasoning part, we could explain the agent's akratic action as the result of a conflict between two parts of the soul each of which acts in accordance with its own ATC judgment. Cf. Walker (1989, 663).

21. The *Republic*'s position on whether one can eliminate all unjust desires is at least complicated and perhaps inconsistent (cf. 559BC, 571B-C, 572B, and Belfiore [1986, 423] and Gill [1985]). It is not clear whether the suggestion made at 485D6-E5 involves internal persuasion and this passage raises important issues which I cannot discuss further here, cf. Gill (1985, 19-21). E.g., does it commit Plato to the difficult idea that all desires have some sort of common 'source'? Does it commit him to the claim that there is at any given time a fixed amount of psychic strength which is apportioned among the agent's desires? Cf. Moline (1981, 77-78).

22. If what makes the former case a proper instance of the Principle of Opposites is that there is a simultaneous desire for drink and a "rejection of the desire" for drink, then it seems the latter should also be a proper instance. Cf. Irwin (1979, 327).

23. *Rep.* 553B8-D4, cf. 442A7-B3 and 429C-430B. Plato does not describe this process in detail, but it seems that a highly anthropomorphized Desiring part will be needed to intervene in the deliberations of the Reasoning part. Cf. the *Laws*' account of "near akratic action" in section 3.

24. Cf. Annas (1981), Crombie (1969), Irwin (1979), Moline (1978, 1981), R. Robinson (1971), Stalley (1975), Williams (1973), and Woods (1987).

25. Here the worry is that the partitioning has illicit explanatory value.

26. For discussions of related issues, see Dennett (1981, 1991), Fodor (1981), and Mele (1987, 80-83 and 142). Both Dennett and Fodor think that we can sometimes make explanatory progress by positing homunculi, but, roughly put, see them as ultimately eliminable in favor of nonintentional explanations.

Unlike Dennett and Fodor, we might be willing to accept ineliminable intentional homunculi, but still think that the Partition Theory lacks explanatory value.

27. Cf. Burnyeat (1976), (1990, 52-61) and Nussbaum (1980, 410-15). If Burnyeat is right about the *Theaetetus*, Plato is concerned in the later dialogues to correct some of his earlier views which undermine the unity of the person.

28. Cf. Dennett (1991), Parfit (1984, 199-306), and Korsgaard (1989).

29. Parfit (1984, 223).

30. E.g., 443C9-444A2, 550A4-B7 and 553B7-D7. Since our ordinary judgments about responsibility assume that there is a single subject or agent to be held responsible, the Partition Theory threatens to undermine these judgments. If Plato holds to the view that the person is a compound of the parts, he will need to explain the relation between judging that the person is responsible and judgments about the parts of the soul. If we instead identify the person with only the Reasoning part, there will be problems about holding the person responsible for undergoing akratic conflict as well as for acting akratically. These issues are especially troubling since in the three passages just cited, the person as something over and above the three parts seems to enter the picture as the entity which is responsible for establishing the proper relations among the three parts. Plato might want to restrict responsibility to the Reasoning part, but he does not face the issue squarely and his views on responsibility are quite problematic. For the *Republic*, see 431A3-B7 and 617D6 ff., and cf. Annas (1982); for the *Laws*, cf. 904A6-C4.

31. Elements of Plato's conceptual framework here—the division of the soul into parts, conceptualizing desire or appetite as, or as involving, a psychic movement or command—recur in ancient psychology, *inter alia* in Aristotle and the Stoics. But in order to determine how closely these theories agree with the *Republic*, we must consider how the apparently shared concepts and claims are embedded in each theory. In particular, theories which divide the soul into 'parts' may only be committed to a division of capacities or powers ( $\deltaύναμεις$ ) or of affections ( $\piάθη$ ) and thus would be quite different from the Partition Theory which divides the soul into distinct agentlike subjects. On the history of partitioning, see Vander Waerdt (1985a, 1985b).

32. Plato in the *Laws* accepts that we can act against both what we believe and what we know to be ATC best: e.g., 689A1-C3, 874E8-875C3 and 902A6-B2. On the distinction between strict and near akratic action, cf. Mele (1987, 7).

33. In the *Laws* (733D7-E3, cf. 731C1-7), Plato continues to hold to his early doctrine (e.g., *Gorg.* 509E) that whenever an agent performs something other than her best option, her action is unwilling ( $\άκων$ ). Thus all strict akratic and near akratic action is unwilling. The precise sense of this claim is unclear without further analysis of what Plato means by 'unwilling,' but since all failure to do what is best is unwilling—including that due to ignorance—it seems that unwilling action is not always the result of physical or psychological compul-

sion. I must leave aside here the relation between Plato's notion of being ἄκων and modern conceptions of nonvoluntary and nonintentional action.

34. For the translation of ἄγωγή as 'pull' or 'drawing' rather than 'cord,' see England (1921, 1: 256).

35. In the *Laws*, 'akrasia' (ἀκράτεια), 'akratic' (ἀκρατής), and 'self-controlled' (ἐγκρατής) are most commonly used with regard to pleasure and pain: 636C6, 710A8, 840C5, 886A9, and 908C2-3. But these are not the only things with regard to which one can have or lack self-restraint: other items (see 869A2-3 and 934A3-6) include fears (φόβους), desires (ἐπιθυμίας), envious emotions (φθόνοις), and anger (θυμοῖς).

36. Nor does the *Laws* share the *Republic*'s view (*Rep.* 440A-441A) that θυμός is by nature an ally of calculation. On θυμός in the *Laws*, see Saunders (1962, 38-41) and T. Robinson (1970, 124-25).

37. 644E1 taken with 645D6-7 shows that Plato here classifies θυμός as a πάθος of the soul and as we shall see below, Plato goes out his way to distinguish clearly between an affection (πάθος) of the soul and a part (μέρος) of the soul (863B3). At 644C6-7, pleasure and pain are described as "two foolish counselors (συμβούλω ἄφρονε)." (Cf. *Tim.* 69D2-3 where boldness (θάρρος) and fear (φόβος) are described as συμβούλω ἄφρονε). But this description is immediately rejected as unclear at 644D4-6 and Plato replaces it with the puppet image. In any case, Plato never in the *Laws* or elsewhere treats pleasure and pain as agentlike parts, i.e., as the subjects of mental states and activities.

38. For related contemporary discussions, cf. Bratman (1987), Mele (1987), Velleman (1989), and Watson (1977).

39. Cf. *Laws* 902A6-B2 and 904A6-C4 and fn. 30.

40. It does not seem to help to posit agentlike parts which are the subjects of the desires and act to increase or diminish their strength, since we are still left with the problem of how the subject's awareness of content leads to a change in strength. Unless we take the unattractive step of attributing to desires themselves a capacity to understand and respond to reasons, we seem to face the same problems we initially did. Thus the postulation of agentlike parts as the real subjects of the desires and beliefs produces no explanatory gain. Indeed, such a move seems to increase the difficulty of explaining rational interaction among beliefs and desires located in different parts of the soul. I am grateful to one of my anonymous referees for pressing this point.

41. *Laws* 649D4-7 suggests that intense pleasure by itself can achieve the sort of effect that drink does; 649C8-D2 and 863B6-9 suggest that pleasure can cause a change in, not just a destruction of, the agent's ATC judgment. Cf. Walsh (1963, 30-33).

42. E.g., *Rep.* 429C5-430C2 and 442B11-C3. Note that the *Republic*'s account, unlike the *Laws*', relies on the idea that the Spirited part can preserve or fail to preserve the right opinion. For Aristotle, cf. *N.E.* VII.3.

43. Neither τὸ λογιστικόν, τὸ θυμοειδές, nor τὸ ἐπιθυμητικόν occurs in the *Phaedrus* and only τὸ ἐπιθυμητικόν occurs in the *Timaeus* (70D7). Nor do the *Phaedrus* and the *Timaeus* refer to the Principle of Opposites. Cf. *Soph.* 230B5-8, which concerns opposed statements and is thus not as general as the *Republic* principle. It is widely accepted that the *Phaedrus* and the *Timaeus* should be dated between the *Republic* and the *Laws*, although it is quite controversial whether they are closer to the former or the latter. (See e.g., Brandwood [1976, xvi-xvii and 1990], Fine [1988, 374-77], Ledger [1989], Nussbaum [1986, 459-60 n. 21 and 470 n. 5], and Vlastos [1983, 27 n. 2]; Reeve [1988, 299-300 n. 56] favors putting the *Phaedrus* before the *Republic*, cf. Guthrie [1987, 396-97]). Fortunately, the question of dating has little effect on my argument. Even if we date the *Phaedrus* and the *Timaeus* relatively late, neither dialogue unequivocally endorses the full Partition Theory of the *Republic* and there is still room for further change in Plato's views. If, on the other hand, we date both relatively early, there is a longer time in which Plato's dissatisfaction with the Partition Theory could have grown.

44. For good discussions of the *Phaedrus*, see Ferrari (1985) and especially Price (1989, 1990). Speaking of separated Forms and the theory of Recollection, C. J. Rowe remarks that the characteristic middle-period ideas found in the *Phaedrus* are found exclusively in the context of a μῆθος and that "in general, the *Phaedrus* seems to expound middle-period ideas only to recommend them for scrutiny" (1986, 121). Cf. Price (1989, 72). On the difficult question of the truth status of μῆθος, see (with references) Rowe (1986, 1989) and Heath (1989a, 1989b).

45. The *Republic* uses εἶδος (e.g., at 439E2) and sometimes γένος (e.g., at 441C6) to refer to parts of the soul. In the *Phaedrus*, there are no uses of γένος and a few marginal uses of εἶδος: 251B7, 253C8, 253D4, cf. 270D5. *Pol.* 309C1-3 distinguishes "the eternal part of the soul" (τὸ ἀειγενές ὃν τῆς ψυχῆς μέρος) from the "mortal" (τὸ ζωογενές) part of the soul. This is a passing reference and does not commit Plato to agentlike subjects of mental states. (Note, however, that true opinions seem to be restricted to the eternal part, 309C-311B.)

46. γένος in τὸ χρουσοῦν γένος (645A7, cf. B1) is simply a highly general word for 'sort' or 'kind' and certainly does not commit Plato to agentlike parts (the same is true for εἶδος at 645A4).

47. *Rep.* 442B5-D1 and 589A6-B6 (cf. 554E3-5) show that this is persuading the Desiring part. Since the agreement includes the Spirited part, it too should be able to make ATC judgments about what is best for itself. λόγος here must mean more than 'speech' and should mean something like 'reasoning' or 'argument,' since suppression by compulsion and fear could involve speech. For the construction with ἀμεινον, see LSJ s. v. ἀμεινον.

48. Cf. Mele (1987, 5-6). *Rep.* 442B5-D1 shows that the Desiring part's judgment, if it is persuaded by the Reasoning part, will be true.

49. Plato seems uncomfortable on this point, since he hedges this denial: “they [the children of the Demiurge] knew that it [the desiring part] would not understand reason [ώς λόγου μὲν οὔτε συνήσειν], and that, even if it did have some share in the perception of reason [εἴ τέ πη καὶ μεταλαμβάνοι τινὸς αὐτῶν αἰσθήσεως], it would have no natural instinct to pay heed to it, but would be especially bewitched both day and night by images and appearances [εἰδώλων καὶ φαντασμάτων]” (*Tim.* 71A3-7). Cf. *Tim.* 70A2-7. I quote, with modification, Bury (1975). I regret that I cannot take up here the interesting issue of the relation between Plato’s views about the intellectual capacities of the lower parts of the soul and those of animals. On animals, see Sorabji (1990).

50. Neither μερίς nor μόριον is ever used by Plato for a ‘part’ of the soul. μερισμόν is found only once in Plato at *Laws* 903B9 which, although interesting, is not relevant to the question of whether human souls have parts. μέρος is commonly used for ‘part’ of the soul in the *Republic* (e.g., 442B11, 442C5, 444B3) and is found in the *Timaeus* (91E6), but not in the *Phaedrus*. There are highly marginal uses of ἴδεα at *Rep.* 588C7, 588D3, *Phaedr.* 237D6, and none in the *Timaeus* or *Laws*.

51. Cf. O’Brien (1957, 87), Rees (1957), and A. Taylor (1960).

52. Could Plato simply have modified the Partition Theory, without abandoning it, so as to meet the various objections seen above? We do, as already noted, see in the *Timaeus* Plato’s dissatisfaction with one important aspect of the *Republic*’s Partition Theory and his resulting modification of the theory. But we have also seen that there are serious problems for any partitioning theory which attributes a significant degree of agency to the parts.

53. The virtues here include all and only courage, justice, moderation, and wisdom. On the Reciprocity of the Virtues and related issues, cf. Ferejohn (1984), Gauthier (1968), Irwin (1979), Penner (1973), C. Taylor (1982), Vlastos (1981, 427-45), (1981a, 1981d), and Woodruff (1976).

54. Since moderation is a necessary condition of justice (e.g., *Laws* 631C7-8), akratics will lack justice; if we construe courage as at 633C8-D3, akratics will also lack courage.

55. I.e., (1\*) Some people know that acting virtuously is ATC best for them, and still often act akratically; (2\*) If a person often acts akratically, then she lacks at least one virtue.

56. The denial of (5) does not entail the Reciprocity of the Virtues since by itself it is, for example, consistent with having either courage or moderation and lacking at least one virtue other than wisdom.

57. Plato may face a similar problem in the *Republic* since in it he accepts the possibility of akratic action, but he does not ever clearly confront the issue there. Nor does the *Republic* employ the ‘broadened’ conception of wisdom that we will find in our analysis of *Laws* 689A1 ff., although the evidence on this point is difficult: see e.g., *Rep.* 428A11-429A3 with 431E4-432B1 and 441C9-10, 442C5-9, 444A1-2, and 444B1-8. At *Soph.* 227D4 ff., Plato distinguishes ignorance (ἀγνοία) and stupidity (ἀμαθία), which are vices due to a lack of

true belief from cowardice (δειλία), immoderation (ἀκολαστία), and injustice (ἀδικία), which are due to akrasia. At *Tim.* 86B1 ff., Plato divides foolishness (ἀνοία) into madness (μανία) and ἀμαθία, both of which involve a failure to have true beliefs. Cf. Hackforth (1946), Irwin (1979, 206-208), and O’Brien (1967, 186-197).

58. σοφία and φρόνησις and their cognates are used interchangeably here: 689D2, 689D4, 689D5, 696C8, and 710A6; ἀνοία is freely used for ἀμαθία at 689B3.

59. 689D4-6 seems to be making the trivially true claim that consonance is a necessary condition of φρόνησις when φρόνησις is construed as a compound of moral knowledge and consonance rather than the claim that one could not acquire moral knowledge if one did not already have properly trained desires. Plato clearly accepts, e.g., at 689B1-3, that one can possess moral knowledge and still have badly trained desires and does not say that such knowledge must have been acquired when one’s desires were properly trained (cf. 874E8-875C3). Of course, badly trained desires may make it more difficult to acquire moral knowledge and may undermine such knowledge once it is possessed.

60. It also, for instance, has important implications for the sort of moral education people should undergo: since internal persuasion is no longer possible, Plato must find other ways of fostering courage and moderation.

61. This is a fairly widespread view, see, e.g., Vlastos (1981a, 1977). Although Plato is more pessimistic in the *Laws*, he may allow for a few exceptions, cf. *Laws* 875D3.

62. Vlastos (1981a, 216).

63. Vlastos (1981a, 216 n. 29).

64. We cannot, of course, know for certain which way the influence ran: Plato’s increased pessimism about the capacity of human beings to avoid akratic action may have led him to rethink his analysis of it.

65. While it seems plausible that self-control will fail in the face of sufficiently powerful desires (cf. 645A2-B1), we still need to know why desires for nonoptimal outcomes will remain and be stronger (both before and after the person intervenes) than the desire for the optimal outcome. In the *Laws*, Plato does recommend educating children to love and hate the right things (see e.g., 653A ff. and 782D-783B), but he does not expect that even in the properly educated the strength of the desire for the best outcome will always be greater than the strength of desires for nonoptimal outcomes.

I would like to thank Alan Code, John Ferrari, Geoffrey Lloyd, Tony Long, Ian Mueller, Martha Nussbaum, Sam Scheffler, Charlotte Stough, Bruce Vermazen, Bernard Williams, Richard Wollheim, and two anonymous referees for their comments on earlier versions of this paper.

### **III. Arguments for Immortality of the Soul**

## The Soul and Immortality in Plato's *Phaedo*

David Bostock

### A. Preliminaries

The discussion of the main topic of the *Phaedo*, i.e., the immortality of the soul, begins at 63e8 with Socrates offering a further "defense." Like the defense that he offered at his trial it is intended to justify the way he has lived, but more particularly it is intended as a justification of his present attitude in the face of death. His main theme is that the true philosopher has been *practicing* for death all his life. As he develops this theme we begin to learn what he takes a "true philosopher" to be, and what he thinks is actually going to happen to him at his death. Throughout this part of the dialogue (63e-69e) the immortality of the soul is simply assumed—of course, the rest of the dialogue will be devoted to proving it—and we may regard the discussion here as mainly an exposition of *what* has to be proved. In particular, I shall concentrate on the question what exactly it is that Socrates calls a "soul" and takes to be immortal.

He begins by defining death as the separation of soul and body, and the state of being dead as the state in which soul and body exist separately from one another (64c4-8). The second part of this definition is somewhat careless, for there is obviously no reason to insist that a man's body must go on existing all the time that he is dead (as is later recognized, e.g., at 80b-c). Those who are skeptical of immortality will evidently say the same of the soul. Setting this aside, another point worth observing is that the definition is somewhat vague about *what* it is that can be said to die or be dead. For the most part Socrates speaks of a *person* dying or being dead, meaning thereby that his soul and body

have separated. But on one occasion he also speaks of a *body* dying (106e6), which we can understand as a matter of that body being separated from its soul, and three times he speaks of a *soul* dying or being dead (77d2-4, 84b2, 88a6), which is presumably to be taken in the same way. Death as separation, then, can easily be understood as applicable to persons and to bodies and to souls. But this last usage is somewhat awkward, for in the Final Argument Socrates will claim that souls "do not admit death." (His position, obviously, is that souls can be separated from bodies, but cannot be destroyed. Perhaps he should have admitted that the word "death" has other meanings than that which is here assigned to it.) But for the present these problems are hardly serious, and I turn instead to the much more fundamental problem that underlies this definition: what is meant by a "soul," and in particular a soul "separated from the body"?

One must of course begin with the observation that the Greek word *ψυχή* (*psyche*) is not very well translated by the English word "soul." I shall continue to use "soul" as a purely conventional translation, but the English word has several unfortunate associations. (For example, it is often reserved for use in religious contexts, which is by no means the case with the Greek word. In more idiomatic speech "soul" is often associated with emotion and feeling, in opposition to reason or intellect, but again there is no hint of this with the Greek.) What, then, does the Greek word mean?

## B. Activities of the Soul

The Greek word has *many* meanings. At its widest it simply connotes life, in a very general sense. For example, the living and the nonliving are contrasted in Greek as the things with soul and without soul ( $\tau\alpha\ \xi\mu\psi\chi\alpha$  and  $\tau\alpha\ \alpha\psi\chi\alpha$ ), and in this dichotomy it is by no means unusual to count plants as things with soul, just because they can reasonably be called living things. Thus Aristotle, who explicitly recognizes a variety of different kinds of soul, has what he calls the "nutritive" soul ( $\tau\o\ \varphi\psi\tau\iota\kappa\o\n$ ) as his lowest kind, and this is the kind of soul that plants have. His point is that they exhibit life by taking in nourishment and so growing, and because this is something that only living things can do there must be a corresponding kind of soul responsible for their doing it. All ordinary<sup>1</sup> living things have this "nutritive" soul, but many will also have other kinds of soul in addition. For example, animals typically have "locomotive" souls, which just means that they also exhibit life by moving around, and so on.

The general principle behind Aristotle's approach is clearly to begin by classifying the different kinds of activity that living things exhibit, and in virtue of which we class them as living, and then to associate a different kind of soul with each. It will be best if we too begin with this line of approach, listing some of the activities that can be credited to souls, so that we can sidestep the awkward question of what kind of a thing the soul *is*, and concentrate rather on what it *does*. (Equally, I shall begin by being somewhat careless over the question whether it is quite right to say that the soul itself *does* this or that, or whether we ought strictly to say that it is the person who does these things, by virtue of the soul that he has. I shall take up later the question of the relation between the person and his soul.)

First, then, we should note that "soul" may be invoked to account for any activity typical of living things, including those more or less biological activities that human beings share with even very "lowly" living things, for example nutrition and growth, reproduction, locomotion, and so on. But without attempting any more detailed list here, I shall pass straight on to what we may call *conscious* activities, for consciousness seems to us a very important ingredient in the typical activities of men (and higher animals), and most distinctive of their lifestyle. Within this general area there are still a wide variety of activities which can usefully be distinguished.

(i) *Perception* is perhaps the most simple form of conscious activity, as when one feels hot or cold, or hears a loud bang, or sees a blue sky. But it should be noted that we also speak of perception when much more sophisticated activities are concerned, for example seeing that the approaching bus is a no. 7. One cannot have this kind of perception without (in this example) some knowledge of what buses are, and of the significance of this way of classifying them; so altogether more complex mental apparatus is involved. Along with simple perceptions one might include feeling pain, which seems just a special kind of perception. In a different direction, many have sought to assimilate *memory* to perception, and one might also wish to add the peculiar phenomena of dreaming, hallucinating, and so forth.

(ii) Perception very often leads to action because we have *goals*, or *desires*, e.g., the avoidance of pain. Desires again range from very simple kinds, such as hunger or thirst, which are quite naturally called "bodily," to much more sophisticated kinds. Consider, for example, the differences between wanting a drink, wanting to watch television, wanting to be prime minister, wanting to know who put the drawing-pin on the chair, wanting to understand differential equations, and so on and on indefinitely. Clearly there is scope for all manner of classifications of different kinds of desire. Along with desires we might mention the

enjoyment or pleasure one may get from satisfying a desire (or in other ways), and the unpleasantness of pain or hunger, or again of continuing to be baffled by differential equations.

(iii) Desires are in some ways similar, and in other ways dissimilar, to *emotions* such as fear, love, hatred, compassion, and so on. Some emotions (such as fear or embarrassment) are typically short-lived states, involving special kinds of feelings, and often with marked physical manifestations (e.g., sweating, blushing). Others are long-lasting states of mind with no such accompaniments, and shade off towards those states of mind or dispositions which we count as part of a man's character. (Compare pride, which is sometimes counted as an emotion, and humility, which usually is not, with, for example, meanness or generosity.) Along with the emotions we might perhaps mention moods such as boredom, hilarity, depression, exhilaration, and so on.

(iv) Most perceptions, desires, and emotions inevitably presuppose *knowledge* of some kind, or at least *belief*. Some beliefs arise fairly directly from perception ("This is water"), some are based on past perception and memory ("Water satisfies thirst"), some are due to reasoning of various kinds ("Drinking-water is colorless. This water is dark green. So it probably isn't fit to drink"). *Reason*, again, may be classified into various forms. In one aspect (just hinted at), reason is often thought of as controlling our desires and emotions, and therefore directing our actions. But in another aspect it may be seen as operating "itself by itself," as when we work out a sum simply because we want to know the answer, and with no intention to act on it. This is often called "theoretical reason," contrasted with "practical reason," and I shall have more to say about it as we proceed, since it is clearly important to the *Phaedo*.

This little catalogue of conscious activities (and states) makes no claim to completeness, but does illustrate the wide variety of forms of consciousness: we are typically conscious of almost all the activities and states just listed, and since the soul is specially connected with consciousness they can all be credited to the soul. (On the face of it, they could almost all occur unconsciously too; whether they should then be attributed to the soul is perhaps a matter for debate, which I shall not enter into.)

### C. The Disembodied Soul (63E-68C, 80C-84B)

When it is claimed that a man's soul is immortal, or anyway that it goes on existing in separation from the body, this claim implies that some at

least of these (or similar) activities continue in the separated state. Perhaps this is not strictly necessary. One might be able to make sense of a theory according to which when the soul was separated from the body it remained *dormant*, and nothing at all happened in it. But clearly Plato's theory is not of this kind. He holds that the separated soul does indeed remain conscious and active. What kinds of activities, then, does he think do continue in the separated state?

When one reads Socrates' Defense (63e-69e), one is inevitably given the impression that the only activities of the soul that can survive separation are the activities of the true philosopher. The philosopher, we find, averts his attention from such pleasures as food and drink and sex (64d4-7), and does not desire material possessions (64d8-e6). Such desires concern the body, and such pleasures come by way of the body, whereas the philosopher has no concern with things bodily. The implication apparently is that since in the separated state there is no body, these desires and pleasures cannot then occur. Next we find that the philosopher despises perception and the use of his bodily senses, eyes and ears and so on (65a9-b7). They are merely a hindrance to what he is interested in, and they too, one presumes, will not occur when there is no body. Similarly, the body distracts him by needing to be looked after, by filling him with irrelevant lusts and desires, and with fears (and presumably other emotions) (66b7-c5). Again the point seems to be that when the body falls away, these desires and emotions will fall away with it, and in the separated state there will be no such things. What the philosopher is concerned with is called "reasoning" ( $\tau\ddot{o}\lambda\gamma\zeta\epsilon\sigma\theta\alpha\iota$ ; 65c2) and described as reaching out for the forms "by pure intellect alone" ( $\varepsilon\iota\lambda\iota\kappa\iota\eta\zeta\delta\iota\alpha\nu\iota\alpha$ ,  $\alpha\dot{u}\tau\eta\kappa\alpha\theta\alpha\iota\alpha\mu\tau\eta\eta$ ; 66a2). Pursuing this goal, the philosopher is already separating his soul from his body as much as he can, and in that way practicing for the complete separation that death will bring. We must presumably infer that pure reasoning will continue into the separated state, but nothing else will.

Before proceeding, I should add a few notes on what we have had so far. First, we are not here offered any very good explanation of why the philosopher despises his senses (65a9-b7). It is claimed that they are not accurate or clear, but no argument is offered for this, and it is anyway not much of a ground for rejecting them as a source of knowledge. But as we read on it becomes clear that the real complaint about them is not that what they tell us is inaccurate, but that what they tell us about—the physical world—is something that is of no interest. For the philosopher's concern is with forms, and forms are not to be grasped by the senses anyway, but by pure intellect. There is no explanation at this point of how one can pursue an intellectual enquiry into anything if one

cannot at least *start* from information supplied by the senses, and indeed this would seem to be a difficult problem. For centuries philosophers have supposed that all knowledge must begin with experience, for what other starting-point is there? But Plato clearly does not agree. He thinks that genuine knowledge is always knowledge of what he calls "forms," and is to be attained by a quite different method, i.e., the "method of hypothesis" (explained mainly at 94d-100a and 101d-e). I cannot discuss these topics here, so I simply accept for present purposes what Plato has to say about forms and our knowledge of them.

A second point worth making here is that Plato is not saying, as a first reading might suggest, that while it is the soul that engages in reasoning it is the *body* that perceives, desires, fears, and so on. If that were his view, it would be difficult to explain how the body's doing these things could so upset the soul, or why getting rid of these desires and emotions should count as purifying the soul. (It would also be difficult to explain the different view of the disembodied soul that we shall meet shortly.) Rather, it is the soul that actually does these things, but it does them when it is in a body, and because of the body it is in. Thus when it sees it sees through the eyes of the body, and will no longer see when it has no body to furnish eyes. Similarly when it wants food it does so because of the feeling which it gets from the body's empty stomach, which again is a feeling it will be free of when it has no body. Generally, while it is in a body it will be sensible of what happens in that body, and this sensibility will inevitably give rise to perceptions, desires, and emotions of all kinds. They happen in the soul, but they would not happen if there were no body for the soul to be sensible of.<sup>2</sup>

This brings me back to the main point that I wish to make. Although there is no direct statement in the *Phaedo* that it is *only* the faculty of pure reason that survives into the disembodied state, that must surely be the implication of Socrates' Defense. The philosopher is described as one who attempts to separate his soul from his body as much as he can, and in effect what this comes to is that he concentrates all his efforts on pure reasoning and pays as little attention as possible to the perceptions, desires, and emotions which arise only because he has a body. This is said to be practicing for death, which is the complete separation of soul from body. It must surely follow that in death the soul has no such perceptions, desires, and emotions; it rejoices always in pure reason and nothing else.

But it is now time to look at some of the *other* activities which are elsewhere credited to disembodied souls. We have not completely described even the state of the dead philosopher, for we have omitted to mention that he will be happy, and will enjoy the society of other gods and (possibly) other men (63b-c, 69e, 81a). But with ordinary nonphi-

losophers our description has been way off target, for it seems that they practice pure reason no more after their deaths than before. Rather, they *fear* Hades (81c11), they *retain* their desires for things bodily (81e1), and they keep their characters as virtuous or vicious, social or antisocial, mild or cruel, temperate or gluttonous, and so forth (81e-82b). If we may add the evidence of the closing myth, they also retain a *memory* of their past lives (108b), and can meaningfully be punished and rewarded (133d-e). Dead souls can *appeal* to one another, *persuade* one another, *forgive* one another, and so on (114a-b). In these passages it appears that disembodied souls are capable of pretty well all those conscious activities that embodied souls are capable of: they can perceive (though presumably without eyes), they can feel pain (though without nerves), they can be frightened (though without adrenalin), etc. I do not wish to imply that this picture is self-contradictory—I do not think it is—but certainly it seems rather less plausible than the more economical picture we had first.

It may be suggested that Plato does offer some explanation of how this could be so: when a nonphilosopher dies, he suggests that the soul is not after all *completely* separated from the body, but remains "interspersed with a corporeal element" (81c4). But this is surely not an explanation that we should take seriously (and its application to ghosts is presumably humorous). If we do take it seriously, then it will imply that the nonphilosopher's soul is in life extended throughout his body, and retains this shape after death, with some material particles somehow "clinging" to it. We can then contrast this with the spatial language which is sometimes used of the philosopher's soul, which is said to "assemble and gather itself together, away from every part of the body, alone by itself" (67c8; cf 80e5). The philosopher's soul is perhaps squeezed into a tight ball, so dense that there are no gaps or chinks where a material particle might be embedded, and so smooth of surface that none can stick to it. The theory can then be tested by accurate weighing of bodies just before and just after death: we expect to find that a nonphilosopher's body suffers some weight loss at death, while a philosopher's body does not. But obviously this interpretation of the *Phaedo* is absurd. It treats the soul as if it were made of some quasi-material stuff, and just the kind of thing that might be blown apart by the wind, especially if you happened to die in a storm (77e1). It is not what Plato means to suggest at all, and when he spoke of a soul being "interspersed with a corporeal element" he obviously meant to be understood as speaking figuratively. His point was just that the soul retains its desires for things bodily.

The belief in a reasonably "full" mental life after death is common, and from Homer onwards (*Odyssey* 11) all those who have pictured it

have pictured the souls of the dead as having the shape of human bodies, and as doing just the kind of things that ordinary living human beings do. Upon reflection one has to admit that this picture is not to be taken literally, but we continue to use it because we do not know any better picture to substitute for it. Plato is no exception. He has no explanation of how disembodied souls can continue with very much the same range of conscious activities as living humans have, but he evidently believes that they do, and so he pictures them as if they were living human beings. We do him an injustice if we think he took the picture seriously.

The *Phaedo*, then, contains two distinct views of life after death. One, which is very much a philosopher's view, and is applied to the philosopher's death, supposes that at death all those aspects of conscious activity which depend upon the soul's awareness of its body will fall away, and as a result the disembodied soul will be capable of pure reasoning but nothing else. The other, which is the more usual religious view, and is applied to other deaths, supposes that pretty well all the conscious activities of ordinary living human beings will persist into the disembodied state. The two are reconciled by a religious doctrine, which Plato took over from Orphics or Pythagoreans, a doctrine of sin, purgatory, reincarnation, and eventual purification and release from "the wheel of rebirth." In more detail, the doctrine is that when you die then you are punished for your sins either by a longish period in purgatory (as in the *Myth*), or by a brisk reincarnation in a less pleasant form of life (as in the elaboration to Socrates' Defense), or by first one and then the other, which seems to be Plato's standard version (as in the myth of the *Republic*). So the object is to live a life that avoids sin—or, if one such life is not enough, then sufficiently many (usually three, according to the myth of the *Phaedrus*). This is "purification" (*κάθαρσις*), and will release you from the wheel of rebirth and admit you to bliss everlasting. To obtain our reconciliation, all we have to do is to adapt this religious doctrine. We take bliss everlasting to be the philosophic afterlife, in which the soul pursues pure reason and nothing else, and we take the life of purification to be the philosophic life, in which the body and all its works are held in contempt, and all energies are concentrated on the life of pure reason. The point here is that pure reason can be pursued in complete independence from all things bodily, and it is therefore the most admirable feature of the soul. (As Aristotle was to stress later, it is also the feature in which human beings most differ from other animals.) It will then follow as a corollary that sin must be equated with paying attention to bodily things. If killing others, injuring others, and behaving violently towards one's parents are to count as sins (114a), that must be because they are cases of paying attention to bodily things, for the scheme now requires that that and nothing else be the criterion of sin.

Now that we have the full doctrine before us, a point that may be observed in parenthesis is that Socrates' Defense is misexpressed. Socrates says simply that he has spent his whole life practicing for death, but he does not mean that he has been practicing for what happens to all men when they die. On the contrary, he has been practicing for a very special kind of death, the death that admits you to bliss everlasting. That is not, however, a very serious criticism. It is more important to consider the implications which this scheme has for morality, implications which Plato himself points to in 68c5-69d2.

#### D. Implications for Morality (68C5-69D2)

Socrates begins with the claim that what is commonly called bravery (*ἀνδρεία*) belongs especially to philosophers, and what is commonly called temperance (*σωφροσύνη*)<sup>3</sup> belongs only to them (68c). Both these claims may be questioned, because Socrates seems to have misidentified what are commonly called bravery and temperance. Bravery he takes to be simply a matter of fearlessness in the face of death (68d), and no doubt it is fair to claim that the philosopher does have this characteristic. But in fact we commonly think that there are many other situations in which bravery may be displayed, e.g., by standing up to physical or mental torture, by being prepared to risk life or health or goods in a noble cause, or simply by being calm or cheerful in conditions of pain or adversity. (Plato had discussed the question "What is bravery?" in the *Laches*, and he there saw a much wider range for it.) There seems to be no special reason why the philosopher should display these other sorts of bravery as well. With temperance the objection is rather different. We can grant Plato that what is commonly called temperance concerns those kinds of desires that he calls "bodily," but ordinarily temperance is taken to require moderation or well-orderedness in the pursuit of these desires (*ἔχειν κοσμίως περὶ αὐτάς*), and not scorning them altogether (*ἔχειν ὀλιγάρως*) (68c8-10). The philosopher, it seems, is a singleminded ascetic who suppresses all his bodily desires, and this is not what is ordinarily called temperance (or moderation, or self-control) at all.<sup>4</sup> (Nor is there any reason to suppose that only philosophers have it: misers and power-seekers may have it too, as our text appears to recognize at 82c5-8.)

Socrates goes on to say that there is something strange and illogical (*ἄτοπον*, 68d3; *ἀδύνατον*, 68e3) about ordinary bravery and ordinary temperance. Ordinarily, he says, a man is "brave through fear"—fear of something he regards as worse than death, perhaps shame, or dishonor.

Similarly the ordinary temperate man is “temperate through being overcome by pleasures,” insofar as he abstains from pursuing some pleasures only because he desires others. (For example, he abstains from drinking too much tonight because he wants the “pleasure” of a clear head tomorrow.) There is actually nothing strange or illogical about this, but it is made to seem strange when the claim that the brave man is brave “through fear” (of greater evils) is re-expressed as the claim that he is brave “through cowardice” (68d12), and similarly when the claim that the temperate man is temperate through his desire for other pleasures is put as his being *overcome* by these other pleasures, and hence as his being temperate “through intemperance.” So apparently these virtues are “caused” by their opposite vices. But both these extensions are illegitimate.

Intemperance is not a matter of being led by *any* desire—e.g., by the desire for health—but rather of giving in to *immoderate* desires, and especially immoderate bodily desires, when one ought not to. Similarly, cowardice is not a matter of being led by *any* fear. Plato himself earlier specified the relevant fear as specifically the fear of death, and on this account being led by fear of shame or dishonor would not be cowardly. But anyway Plato’s simple characterization obviously will not do. Suppose I am thirsty, but abstain from drinking the weedkiller through fear of death, and instead go off to get a drink of water. Surely that is not a case of cowardice? Perhaps a better suggestion might be this: just as intemperance is giving in to *immoderate* desires, so perhaps cowardice is giving in to *unworthy* fears, when one ought not to.<sup>5</sup> The ordinary temperate man, according to Plato, resists his immoderate (bodily) desires only because he has other desires too that he wants to satisfy, but there is no reason why these other desires should also be immoderate bodily desires. Similarly, the ordinary brave man may resist his unworthy fears only through fear of something else, but again that other fear need not be an unworthy one. Some distinctions such as these are certainly required, for otherwise, if we adopt Socrates’ way of talking, every rational action whatever will be done through cowardice and through intemperance. If it is a rational action, there will be some purpose that the agent is trying to achieve, and we can always say that he acts out of a desire for the pleasure of achieving that purpose, and out of a fear that unless he so acts it will not be achieved. This applies even to the philosopher: he wants the pleasure of philosophizing and fears to be deprived of it. But Socrates would not like us to conclude that he is therefore intemperate and cowardly.

Although we can in this way remove the “illogicality” that Plato professed to find in the ordinary man’s behavior, we have not touched the main complaint that he has against it, namely that its motivation is

purely hedonistic: it consists in trading off one pleasure against another, and has as its object just the all-round maximization of pleasure (69a6 ff.). Clearly he is thinking here of more or less bodily pleasures; at any rate he is certainly excepting the pleasure of philosophy. But it would appear that once we add that into the account, the position of the philosopher is not essentially different. He is equally bent on maximizing pleasure, we might say, and his life differs from others only because he enjoys philosophy (or “wisdom”; φρόνησις) much more than anything else. So he obtains his greatest pleasure by putting all his energies into philosophy and paying no heed to the things of the body, but for all that it still seems to be pleasure that he is pursuing. Perhaps, indeed, he is pursuing it more efficiently than others, because he is also taking into account the fact that his way of life will lead swiftly to bliss everlasting, while other ways will not. His hedonism, therefore, takes the longer view, and is very much better thought out. But is it not still hedonism?

Perhaps some light would be shed on this if we could follow out the contrast between “what is commonly called” a virtue and the corresponding “true virtue,” for Plato evidently intends such a contrast, but seems to lose sight of it. One expects him to say that whereas ordinary bravery consists in overcoming the fear of death with the help of some other and countervailing fear, the “true bravery” which is characteristic of the philosopher consists in his not having the fear of death in the first place. Equally the philosopher’s “true temperance” will consist in his not having the temptation to indulge in bodily desires, whereas ordinary temperate men do have such temptations but overcome them (when they are immoderate) with the help of other desires. I think, indeed, that this is what he means to say about true temperance and true bravery at 69b8-c3. Can we extend this idea to explain in the same way the notion of “true justice” (δικαιοσύνη) and the generalization to “all true virtue” (καὶ συλλήβδην ἀληθής ἀρετή; 69b3)?<sup>6</sup> If so, then the thought will be that ordinary men are tempted away from virtue because of their interest in things bodily, and consequently if they do continue to act virtuously it can only be by resisting these temptations with the help of some countervailing “temptation.” By contrast, the true philosopher never is tempted away from virtue in the first place.<sup>7</sup> Plato is claiming, then, that the *only* reason why men are tempted from the path of virtue is that they pay too much attention to the body and all its works. But once a man sees that the only valuable thing is wisdom, and devotes himself singlemindedly to this, then all vicious temptations will fall away and his conduct will automatically be virtuous. That is what counts as “true virtue.” (Admittedly this interpretation goes somewhat beyond anything that is to be found in our text. The very obscure char-

acterization of “true virtue” as a matter of recognizing that wisdom is the only “right coin” (69a9-10) is a decidedly inadequate way of making the point I am here attributing to Plato.)

Supposing that that is Plato’s doctrine, then he is doubly wrong: there are *other* temptations from the path of virtue than those which stem from the body, and the singleminded pursuit of wisdom will *not* preserve one from them. Indeed if we pursue wisdom in so devoted a fashion that this pursuit governs our every action, then there will be many virtues that we will lack. Plato has made out a case for saying that our singleminded philosopher will have a kind of courage, and that he will have a kind of temperance (though his “courage” seems limited to one particular situation, and his “temperance”—which seems more the asceticism of a fanatic—is far removed from the ordinary virtue). But consider now some virtues that he will *not* have. We can, in fact, begin with justice: why should it be supposed that one whose whole ambition is the pursuit of “wisdom” should treat others justly and fairly? How would that help him in his one overriding pursuit? Why, indeed, should he be kind, considerate, loyal, merciful, generous, helpful, forgiving, and so on? How would that assist his intellectual enquiry? The answer must be that it would not. The demands of *other people* are just as unwelcome distractions to him as are the demands of his own body, and he will withdraw from them as much as he can. Yet morality, as we think of it, is primarily a matter of how one behaves towards others. The virtues of courage and temperance are in fact untypical, and are often distinguished as “self-regarding” virtues, since they *can* perfectly well be manifested in actions which do not affect other people. But most virtues are “other-regarding,” and essentially concern one’s behavior to others. These virtues it seems that our philosopher will lack altogether. At any rate the temptation to act only with his own ends in view is a temptation he will certainly have, and apparently he will see no reason to resist it.

In short, the morality which our “true philosopher” lays claim to is thoroughly egocentric. Perhaps it need not be classed as a kind of hedonism. At any rate, our dialogue lays no stress on the *pleasure* to be got from the pursuit of wisdom, though this is something that Plato does stress elsewhere (notably *Republic* 9, 580d-588a). Perhaps we could say that so far as the *Phaedo* is concerned wisdom is to be pursued for its own sake, and irrespective of any enjoyment that its pursuit may bring. But still, the philosopher clearly pursues *his own* wisdom.<sup>8</sup> That is the one thing he wants, and the one thing that will get him where he wants to be, off the cycle of reincarnation and away to bliss everlasting. To this one overriding ambition *everything else* is subordinate, not only the demands of his own body but also all sympathy for others, all concern

for justice, and in short practically everything that we consider important to morality. For this to be at all acceptable as an account of “true virtue,” it would have to be argued that a concern for others is in fact needed as a prerequisite for the efficient pursuit of one’s own intellectual goals, and Plato does indeed try to argue in this way in his *Republic*. I cannot discuss that argument here, but I think it is obvious that the proposition to be argued for is, at least on the face of it, very implausible.

After this digression on morality, let us return to our proper subject, the soul.

## E. The Soul and the Person

We have seen that, so far as the ordinary person is concerned, Plato’s view is that in its disembodied state the soul retains pretty much the same range of conscious activities as the person enjoyed when alive, including many which we naturally think of as due to the happenings in the body. Of course the soul will no longer be performing its function of animating a body—neither such unconscious functions as keeping the heart beating nor such (normally) conscious functions as moving the limbs—but a very good range of its activities will survive. Moreover, the dead soul is thought of as keeping the memory of its life on earth, much the same desires, skills, character, and dispositions, and most of what we regard as contributing to a man’s personality. It is therefore very reasonable to say that we can regard what survives as still *the same soul*, and indeed to identify the soul with the person and count it as the continuing existence of that *same person*. When I die, I do not cease to exist, and what survives still counts as *me*, though now I have lost my body much as in life I might lose a limb.

But what happens when my soul is reincarnated in a new body? In what sense does that new body have the *same* soul as I had? What makes it the same? Certainly no new child will be born already equipped with my memories, my knowledge, my character, and so on. It may be tempting to imagine that the new child has a soul made from the same lump of “soul-stuff” as mine was, in the kind of way in which the same lump of gold may be molded into first one shape and then another. But the difficulty here is that souls are supposed to be *immaterial* things, and the notion of an immaterial stuff does not seem to make much sense. When we conceive the soul as capable of existing in isolation from matter, we are conceiving it as a center of consciousness, and there seems to be no sense in which the newborn child will have the same

consciousness as I now have.

One might perhaps make some headway with the idea of a latent propensity for certain activities. For example, if I am good at mental arithmetic, perhaps my soul will easily (re)acquire this skill when it is next incarnated, though obviously it will not be born with it. We may extend this idea to other features of my personality. Those who believe in reincarnation have often wished to say that memory too is carried over in a latent form, though in this case it will practically always remain latent, and never come to the surface. (There are just a few cases of people who claim to remember a previous life, but they form an infinitesimal proportion of mankind.) Since the idea of a permanently latent memory is somewhat suspect, one might take it as part of the theory of reincarnation that the memories will come to the surface again when the soul is next incarnate, and in that state a soul will enjoy memories of all its previous lives—or, if that seems too overwhelming, at least of its fairly recent lives. (Memories may fade with time, but for incarnate souls the timespan will be very long.) However, it must be admitted that this is not how Plato himself seems to envisage it.

The reason why memory seems so important in this issue is that nothing else seems adequate to ensure the *identity* of an immaterial center of consciousness over time. A soul at one time may surely have all the same ambitions, skills, character, and personality as some *other* soul did earlier, and the two need not on this account be the same soul. Unfortunately, the same may be argued to be true of memory as well: though it does not in fact seem to happen, there is surely nothing *impossible* in the suggestion that two souls which are indeed different souls should nevertheless each seem to remember doing exactly the same things. You may reply that only one of them can be genuinely remembering, and the other must be suffering from a delusion of some kind, but then you are faced with the difficulty of distinguishing between genuine and apparent memory in a way which does not presuppose that we can already attach some content to the idea of being the same soul. Since the issues here become very complicated, I shall have to leave this question unresolved.

For the sake of argument, let us suppose that the question is satisfactorily settled, and that we can make adequate sense of the idea that the same soul may occupy now one body and now another, with periods of incarnate existence in between. I shall leave aside the further question whether we have any reason to think that this theory is true, or whether on the contrary it has been upset by facts which Plato knew nothing of (such as the age of the earth and the growth of population). For there is still a further philosophical question which needs raising, namely as to the relation between the soul and the person. On Plato's theory, an

ordinary living person is a combination of soul and body, and at his death that combination is destroyed. Must not that be the destruction of him, the person? Perhaps, indeed, his soul will live on earth again in another body, but that is surely not the same as to say that *he* will live again. I think it would be generally agreed that if, when I die, your brain is transplanted into my body and the heart is started up again and the body made to live once more, then that may be another life for that body, but it will not be another life for *me*. Assuming the liaison between brain and consciousness that we normally do assume, the resulting person would have your consciousness and my body. But I am not to be identified with my body, and though the body may live again it does not follow that I shall. Why should we not say exactly the same thing about the soul?

The answer is implicit in the example I have just introduced. If your brain is transplanted into my body, and as a result your memories, your mental capacities, your character and personality change from one body to another, then indeed the resulting person is not I. But surely it *is* you? We are familiar with the idea that you might have a kidney transplant or a heart transplant, and this seems merely to be the extreme case of a whole-body transplant. As one might say, where your consciousness is, there also are you. Generally, if a person is a combination of a soul, conceived as a center of consciousness, and a body, then the soul is the dominant partner so far as the identity of that person is concerned. Admittedly, we have seen that there are problems in understanding the notion of "same consciousness," problems which are partly sidestepped in my example by bringing in the brain, and assuming that if we have the same brain then we shall have the same consciousness. But if these problems can be favorably resolved, then it does not seem unreasonable to identify the person with his consciousness.

But now notice that it is crucial to this theory that being the same soul should be a matter of continuity of consciousness, for if we adopt any other account of being the same soul it will no longer be reasonable to identify the person with his soul. For example, suppose we think of a soul merely as a kind of animating agent that makes a body live, and not specially connected with consciousness, much as an engine may "animate" a car. The same engine may be transferred from one car to another—cars of very different shapes and sizes, and very different performance on the road—and there is no temptation to say that these cars are all really "the same car," just because they have the same engine. (On the contrary, a car may be given a new engine and still remain the same car.) Similarly for souls. If what now animates me will one day animate an astronaut, but there is no continuity of consciousness between me and that astronaut, then it is quite unreasonable to regard his

life as another life for *me*. I am not my “animating agent” on this view, and *I* shall never be an astronaut.

With these points in mind, let us consider Plato’s own version of the theory of reincarnation. When I die, then since I am an ordinary person who has not spent his life withdrawing from all things bodily, it will be reasonable to say that I do survive into the disembodied state. For in the case of ordinary persons Plato does provide as much continuity of consciousness as one could ask between the embodied and the disembodied soul. But the difficulty comes when we consider the next transition back to the embodied state, for here Plato seems to provide no continuity of consciousness at all. Indeed he envisages my soul returning to earth to inhabit a donkey, a hawk, or an ant, and it is impossible to see how the consciousness of such a creature could still be *my* consciousness. (When Apuleius told how he was transformed into a donkey, he convinced us that the donkey was indeed he by allowing it to retain his memories, his desires, and his personality. That made it a very unusual donkey.) Perhaps one can find some sense in which it is still my soul that will be animating these later creatures, but in that case I am not my soul, and though my soul may be immortal, I am not.

In the case of ordinary people, then, it is the transition from the disembodied state back to the next embodied state which creates the greatest problem, but this problem at least does not arise for the “true philosopher.” When he dies, his soul never does return for a further earthly life. But here we have another problem: does *enough* of his soul survive death for it to be reasonable to say that *he* survives? At a first glance it appears that the answer must be “no,” for the only thing about him that persists beyond his death is his capacity for abstract reasoning, reasoning which does not in any way depend on what he has learnt through his body. This appears to be so small a part of his total personality that it is unreasonable to identify it with him, and difficult to say in what sense it can even be identified with his soul. Are we to suppose that at least his soul will remember its experiences on earth, even though it has no use for such memories? Since Plato gives us no information on such points, it seems rash to speculate. But perhaps we can just say this: during his life the philosopher has as it were “identified himself” with his capacity for pure reasoning, insofar as that is the only thing he has been interested in, and that may make it seem reasonable to say that *he* survives. On the other hand, my own consciousness of myself as distinct from others seems not to play any role in the kind of abstract reasoning that Plato envisages, so perhaps it should be regarded as one of those things that the philosopher loses at death. But in that case what survives is not even self-conscious, and if so then it surely cannot be identified with the person it has survived from.

Of course Plato speaks, throughout the *Phaedo*, in terms of personal survival. Socrates says “Cheer up! This is not the end of *me*. I am going from here to a better place.” It would surely be less comforting if he had said “This is, after all, the end of *me*, but some bits of *me* will still survive: my bones will last for a little while yet, and my reasoning capacity will be disporting itself elsewhere.” But although it is clearly a personal survival that Plato envisages, we can only say that he has not really seen the problems that this involves.

Since Plato seems never to have abandoned the doctrine of reincarnation, he can never have seen the problem of personal survival properly. But it is connected, as we observed, with the question *how much* of the soul is supposed to survive death, and on this he certainly did have further thoughts. Indeed the *Republic* poses the question more sharply by explicitly distinguishing three different parts which together make up the whole soul. The “top” part is the reasoning part ( $\tau\ddot{o}\lambda\gamma\iota\sigma\tau\iota\kappa\acute{o}\nu$ ), and this both pursues theoretical reasoning (about forms) and has the job of controlling the desires. The “bottom” part is called the desiring part ( $\tau\ddot{o}\acute{e}\pi\iota\vartheta\mu\gamma\tau\iota\kappa\acute{o}\nu$ ), and this is responsible for the bodily desires and for such longer-term desires as spring from the bodily desires, such as the love of money. In the *Republic* there is less emphasis on the role of the body in prompting these desires, and more emphasis on the point that the desires are themselves desires of the soul. Finally, the “middle” part is called the spirited part ( $\vartheta\mu\o\iota\vartheta\acute{e}\varsigma$ ), and to some extent it seems to represent emotions, such as anger or self-reproach, while to some extent it seems to be the seat of another kind of desire, the desire for honor and glory and success in public life. It is accounted a “natural ally” of the reasoning part. There are some problems with the precise delimitation of the functions of these three parts, which I do not intend to go into, but one must applaud Plato’s recognition that the soul is not after all such a simple and unitary thing as the *Phaedo* seems sometimes to suggest.

Like the *Phaedo*, the *Republic* also argues that the soul is immortal, but seems curiously undecided as to how much of it is immortal. The argument itself (which is quite unlike any argument in the *Phaedo*) speaks simply of “the soul,” and appears to apply to the whole soul (608d-610e). But then there is a passage—reminiscent of the Affinity Argument in the *Phaedo*—which suggests that since the soul is immortal it cannot really be composite, but must merely appear composite because of its association with the body (611b-612a). However, the question is not further explored there, and it seems that Plato continued to remain uncertain over it. The three parts of the soul reappear in the *Phaedrus* (likened to a charioteer controlling a white horse and a black

horse), and in the *Timaeus* (located in the head, the breast, and the belly, respectively). In the *Phaedrus* it certainly seems to be implied that the compound of all three parts is immortal, but in the *Timaeus* we are clearly told that only the reasoning part is immortal.

Viewed in terms of the later doctrine of parts, the *Phaedo*'s position is that only the reasoning part is, strictly speaking, immortal. The other parts will standardly survive from one incarnation to the next, but they can be made to wither away by living a suitably philosophic life, and when that happens the reasoning part is freed from the wheel of rebirth and never comes back to earth again. (But then the question arises: how did the soul acquire those other unwanted parts in the first place? Perhaps the *Phaedrus* reflects Plato's uneasiness over this question.) Ideally, we should of course hope to determine how much of the soul is immortal by seeing what is *proved* to be immortal by the arguments for immortality, and this brings me at last to a consideration of those arguments.

## F. The Arguments for Immortality

There are essentially four arguments, namely the Cyclical Argument (69e-72d), the Recollection Argument (72e-77d), the Affinity Argument (77e-80b), and the Final Argument (95e-106e). But between the third and the fourth there comes an Interlude (84c-95c), which is highly relevant to our topic. I take them in order.

The Cyclical Argument aims to show that the cycle of death and rebirth must go on forever, since otherwise everything would end up dead, which is assumed to be impossible (72a-d). (This is quite a curious assumption, from our point of view. It is also at odds with Plato's own view that philosophers can after all escape the cycle of death and rebirth, for why should not every soul end up as a philosopher?) It is clear that what Plato is thinking is that souls must oscillate perpetually between the state of being "alive" (i.e., conjoined with a body) and "dead" (i.e., separated from a body), but this by itself tells us very little about what kind of thing a soul is supposed to be.

The Recollection Argument is perhaps the most interesting argument in the book, and I cannot discuss it properly here.<sup>9</sup> But one can certainly say at least this: it aims to show that some *understanding* which we have in this life can be explained only by supposing that we did not acquire it in this life, but were born with it. The relevant understanding is understanding of forms, and the argument presumes that we can be born with this understanding only if in a previous existence our

souls enjoyed something like a "direct acquaintance" with the forms, and now (dimly) remember it. This does tell us something about the soul: it is where understanding resides, and it has memory. We may fairly generalize this to the claim that what Plato later calls "the reasoning part" must have had a previous existence, but it tells us nothing about the previous existence of desires or emotions.<sup>10</sup>

A similar conclusion follows from the Affinity Argument, which claims that the soul *more* resembles the invisible, unchanging, and eternal forms than it does the visible, changing, and perishable objects that we perceive in this world. The support offered for this (implausible) claim is that the soul is at rest when contemplating forms, but "wanders" and "is confused" when perceiving the objects of this world (79c-e). So again it is the "affinity" between souls and forms that is being relied on, which presumably applies only to the reasoning part of the soul. But Plato further adds a comparison between souls and the immortal gods, on the ground that in each case their nature is to "rule," where the thought is that the soul does—or anyway should?—rule the body (80a). In Aristotle's terms, it is "practical reason" rather than "theoretical reason" that is here in question, so there has been something of a shift. But our present passage says little about the nature of this "ruling," and for that we must look rather to the ensuing Interlude.

In this Interlude the defects of the two preceding arguments are clearly recognized:<sup>11</sup> the Recollection Argument shows only that the soul existed before birth, but not that it will also exist after death, and the Affinity Argument draws analogies in one direction, but there are equally good analogies to be drawn in the opposite direction (84c-88b). More relevant from our point of view is that the suggestion is raised that the soul is not the *kind* of thing that could exist without a body, for Simmias proposes that the soul should be regarded as a "harmony" (or "attunement") of the physical elements of the body (85e-86d). Against this Plato uses once more the idea that the soul *rules* the body, and for that reason cannot be just a "harmony" of its elements (92c-93a, 94b-95a). But in the development of this idea it becomes clear that what he is mainly thinking of is the point that (practical) reason rules (or should rule?) the desires and emotions. So if this is to be an objection to Simmias' original proposal, we must take it that he is here construing desires and emotions as physical elements from which the body is composed. Taken literally, this is simply grotesque. More charitably we may say that the tendency of this argument is to associate desires and emotions very strongly with the body, and to contrast them with (practical) reasoning, which is to be construed as due to the soul and not the body.

Yet Plato interweaves with this argument another, which appears to tend in the opposite direction (93a-94b). This begins with the thought

that a soul cannot be a harmony, since being a soul is not a matter of degree, whereas being a harmony is. But it goes on to suggest that souls differ from one another according to their *own* (internal) harmony, a good soul being more harmonious and a bad soul less harmonious. One therefore asks: what distinct elements are there *in the soul* between which this internal harmony may or may not subsist? Our text can hardly be said to be explicit on the point, but if we may borrow from the *Republic* the answer seems clear: it is the three parts of the soul (roughly, desire, emotion, and reason) that may or may not be “in harmony” with one another. So one may fairly protest that one of these objections to Simmias’ proposal treats desires and emotions as “parts” of the body, while the other treats them as parts of the soul. One can only conclude that while Plato may be on the way to his mature theory, he has not yet got a clear view of it. But in any case these objections have no clear implications for immortality, which we must now return to.

The Final Argument begins with a long and interesting preamble, which focuses on the notion of a cause (or explanation), proposes various conditions on what can properly be counted as a cause, and concludes that the simplest and most straightforward examples are the forms. Thus the cause of a thing’s being or becoming *F* can always be given as (its participation in) the form of *F*-ness (95e-102a). One reason why this kind of cause is especially satisfactory is that the form of *F*-ness is itself a thing that is *F*, and one that cannot be or become not-*F*, and this is the point that is mainly relevant to the argument. For it is then claimed that there are also other examples of causes which satisfy all the conditions (102b-105c), and this is applied to the idea that the soul is the cause of life (105c-106c). The conclusion drawn is that the soul, as cause of life, must itself be alive, and cannot be or become not alive, which is apparently to say that it is immortal. (This conclusion does not in fact follow; all that follows is that a soul cannot both exist and be not alive, but that does not mean that it cannot cease to exist.)

Unfortunately this argument, which Plato eventually (in this dialogue) relies on, quite upsets the view of the soul which has been prevalent hitherto. For amongst his conditions for being a proper cause he has claimed that causes must be both necessary and sufficient for their effects, and so we must suppose that this too applies to the soul as cause of life. Thus whatever has a soul is alive, and whatever is alive has a soul, and this holds not only for man, and for the donkeys, hawks, and ants previously mentioned, but also for things which surely have no “reasoning part,” e.g., aphids, jellyfish, oak trees, cabbages. By this argument, all of these will have immortal souls, so it cannot be essential to

an immortal soul either to be capable of reasoning, or to have desires and emotions, or even to be conscious.<sup>12</sup> One presumes that Plato failed to notice this point. If he had done, he might have started from the premiss, not that the soul is the cause of life, but that it is the cause of consciousness, or of reasoning, or something else of this sort. That would certainly have clarified his conception of the soul (but it would not have improved the logic of his argument).

## Notes

1. For Aristotle, the stars are alive, and they have “locomotive” souls but no “nutritive” souls. God is a living thing with neither of these kinds of soul.

2. In 94b-d it does appear to be the *body* which has desires and emotions and not the soul, but the passage is in any case a surprising one. I add a brief comment on it in my final section.

3. σωφροσύνη is a notoriously untranslatable word. It combines the notions of temperance, moderation, and self-control, and often shades off into meaning just “good sense.”

4. Aristotle would regard it as a *vice* opposed to temperance, which he calls ἀναισθησία (EN 1107b6-8).

5. But are courage and cowardice always concerned with fears? Neither the child who cries when he is mildly hurt, nor the child who bites back his tears when he is badly hurt, need be afraid of anything.

6. Gallop translates “and, in short, true goodness,” using “goodness” (instead of the more usual “virtue”) for ἀρετή. Evidently, a generalization to *all* true goodness-or-virtue is intended (*Phaedo*, ed. D. Gallop [Oxford: Clarendon, 1975].)

7. In Aristotle’s language ordinary men can at best attain “self-control” (ἐγκράτεια), which Aristotle ranks below “virtue.”

8. One could imagine a modification of Plato’s doctrine, in which the pursuit of wisdom was recognized to be a cooperative endeavor, and the object was to make as much wisdom as possible available to the human race. There would still be objections to be raised to this view, though they would not be quite the same objections. But anyway, this would be quite a drastic modification to the religious doctrine of sin, purification, and release from the wheel of rebirth, which was our starting point. That religious doctrine was always egocentric.

9. For discussion of this argument, see Dominic Scott, “Platonic Recollection,” in *Plato 1: Metaphysics and Epistemology*, ed. Gail Fine (Oxford: Oxford University Press, 1999), 93-124.

10. Compare *Phaedrus* (249b-c), which says that no soul can be reborn as a human soul unless it has seen the forms, and so can understand language.

11. No defect in the Cyclical Argument is pointed to, and in fact there is no further mention of this argument (here or hereafter). It is natural to infer that it

drops out of consideration because Plato was not fully satisfied with it.

12. Cf. *Phaedrus* (249b-c), cited above, n. 10. But what, in that case, corresponds to the charioteer and his two horses, of which *every* immortal soul consists?

## 11

## Plato's "Cyclical" Argument Recycled<sup>1</sup>

David Gallop

In his generous review of my *Phaedo*<sup>2</sup> Jonathan Barnes has put forward a reinterpretation of Plato's argument at 69e6-72e2, and an assessment of it which is at variance in certain respects with my own. Since the so-called "Cyclical Argument" is of intrinsic philosophical interest, and raises some wider issues of Platonic exegesis, I would like to make his analysis the focus of some further remarks about it. I begin (I) with some comments on the concept of immortality bearing upon the scope of the Argument. Next (II) I consider Barnes's account of the relation between the Cyclical and Recollection Arguments. I then (III) discuss the major principles of his interpretation. Finally (IV), I reconsider how seriously the Argument should be taken in the wider context of the dialogue.

## I

What does "immortality" mean, and is the Cyclical Argument designed to prove it? Barnes holds that "the Cyclical Argument is not, and was not intended to be, an argument for immortality" (401). It "pretends to show only that souls exist at some time when their owners do not" (401). For its conclusions, as variously stated in the course of the Argument, are compatible with the soul's having only a finite duration: "The possibility of a temporary survival of death is not overlooked in the *Phaedo* (see 87d-88b); and there is no reason to suppose that it is ignored at 69e-72c" (401).

The suggestion that the Argument was not intended to prove full-blown immortality goes back, as Barnes notes (401, n. 12), to Olympiodorus. It was revived some years ago by D. O'Brien,<sup>3</sup> who distinguished between what he called “partial” and “full” immortality, the former being the thesis that the soul can survive separation from the body only a finite number of times, the latter that it can survive any number of such separations. That distinction, O'Brien suggested, would help to explain the addition of the predicate “imperishable” ( $\alpha\nu\omega\lambda\epsilon\theta\rho\varsigma$ ) to what Socrates must later prove about the soul in response to Cebes' objections (88b5-6, cf. 95cl). Mere partial immortality would be insufficient. Only a proof of the soul's “imperishability” would secure the full immortality that Cebes demands.

That suggestion does not, however, accord with the standard usage of  $\alpha\theta\alpha\nu\alpha\tau\varsigma$  and its cognates. Like the English “immortal” the Greek word connotes permanent immunity from death. The Homeric immortals ( $\alpha\theta\alpha\nu\alpha\tau\varsigma$ ) are “those who always are” ( $\alpha\iota\epsilon\nu\ \dot{\epsilon}\delta\omega\nu\tau\epsilon\varsigma$ ). The King of Persia's “immortals” are a guard whose numbers are never increased or depleted (Herodotus VII. 83).<sup>4</sup> In Plato's usage likewise “the immortal” is associated with permanence, often with “the whole of time.” Thus at *Meno* 81b3-6 “immortality” connotes the soul's persistence through an endless series of incarnations. In the *Republic* (608c9-d1) Socrates asks whether an immortal thing “should take thought for a time no longer than the span of life and not *for all time*”; and at 611a2 he derives the soul's immortality from its “existing always” ( $\epsilon\iota\ \delta'\ \alpha\iota\epsilon\ \delta\nu$ ,  $\alpha\theta\alpha\nu\alpha\tau\varsigma$ ). In the *Phaedrus* (245c-e) its immortality is connected with its capacity for omnitemporal self-motion. Elsewhere in the *Phaedo* it has similar associations: at 80b1 and 81a5 it is closely linked with “the divine”; at 80d2 bones and sinews are said to be “practically immortal,” meaning that they last almost forever; and at 106d3 it is argued that the immortal, “being everlasting” ( $\alpha\iota\delta\iota\varsigma\ \delta\nu$ ), must be imperishable. In none of these places does  $\alpha\theta\alpha\nu\alpha\tau\varsigma$  mean possessing merely “partial” immortality or the capacity for temporary survival.<sup>5</sup>

Consequently, if the Cyclical Argument purports to prove the soul  $\alpha\theta\alpha\nu\alpha\tau\varsigma$ , its pretensions cannot be limited to showing only that “souls exist at some time when their owners do not.” But that its purport is precisely to prove the soul  $\alpha\theta\alpha\nu\alpha\tau\varsigma$  seems evident from 73a2, where Cebes says, introducing the Recollection Argument, “in this way too, it appears that the soul is something immortal.” The words “in this way too” ( $\kappa\alpha\iota\ \tau\alpha\iota\tau\eta$ ), which point forward to the coming argument, surely imply that the preceding one was also a way of proving the soul “immortal.” Of

course, it might still be suggested that  $\alpha\theta\alpha\nu\alpha\tau\varsigma$  is being used here atypically to mean merely “partial” immortality. But this seems most unlikely. As we have seen, there appears to be no other place in Plato where its meaning is thus limited.<sup>6</sup> It is also highly significant that later on, when summarizing Cebes' objections to the earlier arguments, Socrates makes him insist that the soul's prenatal existence would not prove its “immortality” ( $\alpha\theta\alpha\nu\alpha\sigma\iota\alpha$ ), *even if the soul were reincarnated many times* (95d5). Here at least there is clearly no restriction whatever upon the meaning of “immortality.”<sup>7</sup>

Nor should we expect the Cyclical Argument to contain any. It is, after all, an ingenious retelling of an “old story” ( $\pi\alpha\lambda\alpha\iota\delta\varsigma\ \lambda\o\gamma\varsigma$ , 70c5-6), a philosophical version of a myth ( $\delta\iota\alpha\mu\theta\iota\lambda\o\gamma\wedge\mu\epsilon\varsigma$ , 70b5-6). But in Greek eschatological myths discarnate souls do not typically enjoy mere temporary survival. Either they undergo purgation through a series of punishments and reincarnations, or they are consigned to endless torment, like the incurables in the *Gorgias* (525c-e) and the *Republic* (615c-616a). So if Plato intended a logical reconstruction of a familiar myth, but adopted a concept of partial immortality, his version of the story would be an awkward misfit.

It remains true, as Barnes says, that the possibility of a temporary survival of death is not overlooked in the dialogue. This idea is certainly present in Cebes' image of the weaver who outlives many of his own cloaks, yet fails to survive the last one (87d-88b). But the Cyclical Argument can hardly be supposed to anticipate that objection. It cannot, that is, be supposed to acknowledge that the soul might perish in one of its separations from the body, as Cebes imagines (88a10-b2).<sup>8</sup> For if it did acknowledge that, Cebes' objection would not be the forceful counterargument that it actually is.

I conclude that there is no good reason to read the Cyclical Argument in the restricted manner proposed by Olympiodorus and Barnes. Whatever its logical merits, the Argument at least purports to prove that the soul is immortal in the full sense, and not merely that it can exist temporarily in a disembodied state.

## II

Barnes distinguishes (400) three different demonstranda for the Cyclical Argument formulated in different places: (A) “our souls exist in Hades after our deaths, and have power and wisdom” (70b2-4); (B) “our souls

exist in Hades after our deaths" (70c4-5), and (C) "our souls exist in Hades" (71e2).

He says that (A) entails (B), and (B) entails (C), but that the converse entailments do not hold. Some commentators, as he says, have thought that the Cyclical and Recollection Arguments are concerned, respectively, with (B) and with the part of (A) referring to "power and wisdom." On the other hand, commentators have disregarded the distinction between (B) and (C). Barnes, for his part, regards (B) as an elliptical form of (A), and thus as needing no supplementation from the Recollection Argument. For "Plato nowhere hints that the power and wisdom of souls in Hades requires a proof distinct from the proof that souls exist in Hades: nor should he have done; for how could there exist a soul which had no power or wisdom?" (404).

Barnes also insists upon an important distinction between (B) and (C), interpreting (C) as restricted, within the Cyclical Argument, to *prenatal* death and existence: "The Cyclical Argument, we might say, concerns itself exclusively with *prenatal* death and existence; and from that nothing follows about *post mortem* death and existence. In short the Cyclical Argument does not prove proposition (B)" (417). He thinks, however, that Plato certainly desired to prove (B), and indeed needed to do so, since (C) by itself is devoid of "eschatological interest" (417). He also suspects that Plato thought he had proved (B), though he did not, in fact, succeed. He has at best proved the much feebler (C), although even there the Argument has to be helped out at certain weak points. As far as eschatology is concerned, the confusion between (B) and (C) is the Argument's "ultimate undoing" (416-17). Moreover, "if the Cyclical Argument is intended to establish (B), it is intended to establish (A)" (400-401), of which (B) is an elliptical form. So presumably, since Plato has not really succeeded in proving even (C), *a fortiori* he has not really proved (B), and *a fortissimo* he has not proved (A) either.

All of this seems to me an unlikely division of labor between the Cyclical and Recollection Arguments. First, if (B) were merely elliptical for (A), then the Recollection Argument would appear otiose. But that is hard to credit. It is true that the Cyclical Argument, if concerned only to prove (B), would need no supplementation for that purpose alone from the Recollection Argument.<sup>9</sup> But if Plato himself viewed (B) as elliptical for (A), why did he offer an argument which reads as if it were intended to establish part of (A) independently?<sup>10</sup> Why did he think it necessary to argue separately for a component of (A) at all?

More fundamentally, it seems perverse to insist upon a distinction between (B) and (C), when the only difference is the omission in (C) of the words "after our deaths" or "when we have died," and yet to view (B) as elliptical for (A), when the former lacks the very significant words "and have power and wisdom." Would it not be more natural to read "our souls exist in Hades" as elliptical for "our souls exist in Hades after our deaths," than to take this latter sentence as elliptical for the whole of (A)? Since Hades was the House of the Dead, one might think that a soul's existence in Hades was posthumous "by definition." When Socrates speaks the language of traditional religion (69c, 80d, 81c, 107d), he talks of souls "travelling" to Hades or "arriving" there. But were there any travellers to Hades who had not once lived in the world above? Were there any souls in Hades of people who had not died, apart from temporary visitors like Orpheus and Odysseus?<sup>11</sup> If not, the words "after our deaths" in (B) are pleonastic, simply expressing something already given in the very idea of being in Hades.

On the other hand, it seems far from obvious that a soul's mere existence entails its possession of power and wisdom. "How," asks Barnes, "could there exist a soul which had no power or wisdom?" The implied answer is, presumably, that there could not exist a soul having neither power nor wisdom. But this need not preclude Plato from treating the possession by souls in *Hades* of power and wisdom as needing separate proof. Certainly, it is clear in the *Phaedo* itself that some souls do not possess wisdom, either in this life (93b8-c1, 94b4-5) or after death (108a-c). And given the wide range of functions ascribed to the soul in different parts of the dialogue, one might reasonably expect some independent proof of the discarnate soul's possessing wisdom as well as power.

This expectation is surely justified, if we consider the traditional Homeric conception of souls in Hades, to which Plato stands in conscious opposition. The souls of the dead visited by Odysseus in *Odyssey* XI, so far from possessing power and wisdom, could not even recognize their visitor, or speak to him, till they had taken a draught of sheep's blood. They bemoan their own weakness, reduced as they are to shadows of their former selves, feeble, dreamlike phantoms with no vigor in them (XI.393-4).<sup>12</sup> "How dared you come down to Hades, where dwell the witless dead, images of toiling mortals?" asks Achilles (XI.475-6). From this picture the idea that souls in Hades possess power and wisdom would constitute a very significant departure. Earlier in the *Phaedo* (67-68) much has been made of the idea that the true philosopher can look

forward to attaining wisdom upon entering Hades. Later on (80d) the conception of Hades will again be transformed. For it is “Hades in the true sense,” viz. the unseen domain of Forms, which Socrates hopes to enter himself.

It therefore seems far more likely that Barnes’s proposition (C) is an ellipsis for (B) than that (B) is an ellipsis for (A). We should not, I believe, regard (B) and (C) as different conclusions at all. As for (A) and (B), they are genuinely different, and the “wisdom” component of (A) certainly does require separate proof rather than subsumption under the Cyclical Argument. It is plausible to find that proof in the Recollection Argument. If that is correct, the Cyclical Argument is not a self-contained argument for (A), but needs the Recollection Argument to complement it. It is a self-contained argument for (B) or (C), although flawed in ways that will appear below.

### III

According to Barnes’s analysis, the Cyclical Argument depends upon two Principles of Change, which he calls (1) the Principle of Opposites and (2) the Principle of Existence. Let us consider each of these in turn.

(1) The Principle of Opposites says that (i) if  $F$  and  $G$  are opposite properties, then any object that acquires the property  $F$  at a time  $t$  possesses the property  $G$  during a period immediately prior to  $t$  and (ii) conversely.

In this formulation “opposites” are to be taken, in effect, as properties that are both mutually exclusive and jointly exhaustive (402). They cannot both be true of a subject capable of possessing them at the same time, nor can they both be false. Any subject capable of possessing either of them must possess one or the other and cannot possess both simultaneously.

One trouble with this account of opposites, as Barnes recognizes, is that it fails to fit all of Plato’s examples, including the very first “opposites” he mentions, beautiful and ugly, just and unjust (70c3). Since there are properties “in between” these, their presence in the text seems extraordinary (as Barnes says, “hard to explain or to excuse,” 405). But we can hardly write them off as a mere aberration on Plato’s part, for they occur as prime examples of “opposites” elsewhere, as do others that would cause similar problems, such as “large” and “small” (102d-e), or “hot” and “cold” (103d-e).

Moreover, with this interpretation of “opposites,” it seems doubtful whether “alive” and “dead” are opposites in the relevant sense, since they are not jointly exhaustive. There remains the option “neither dead nor alive,” applicable to someone who is not yet living. To be sure, we can force “alive” and “dead” into the mold of exhaustive opposites, if we interpret “dead” as meaning “not living” or “inanimate.” But is this what Plato means by “the dead” or “that which is dead” in this argument? Barnes allows (409-10) that neither  $\tau\epsilon\theta\eta\chi\omega\varsigma$  nor  $\alpha\pi\theta\theta\alpha\varsigma\omega\varsigma$  can bear the sense of  $\alpha\psi\chi\omega\varsigma$ ; and it does not seem at all plausible that when Plato writes that the living come to be from the dead, he means that they come from the inanimate. Furthermore, if we do take “dead” in this way, we weaken the argument at a crucial point, namely, the transition from “those who are dead” to “those who have died.” Barnes calls attention to this himself, when he distinguishes (417-18) between a proof that we exist when we *are dead* from a proof that we exist when we *have died*. He says that given the interpretation of “dead” which the Argument requires, “Socrates is dead” does not imply that “Socrates has died.” For this latter implies that Socrates used to be alive, whereas “Socrates is dead” does not. He suggests, as we have seen, that “the argument concerns itself exclusively with prenatal death and existence and that nothing follows about *post mortem* death and existence” (417).

With our ordinary concept of death, however, the idea of “prenatal death” makes no sense at all. Queen Anne could not be dead before she was even conceived. Before that point she was not even alive, let alone dead. Once she is dead, it follows that she has died, and has, in consequence, already lived. Plato’s careful shifting of tenses (at 70c5, c8, c9, d4) trades on precisely these conceptual points.<sup>13</sup> For the entailments between “is dead” and “has died,” and between “has died” and “used to be alive,” hold in Greek just as they do in English.

In my notes (107-108) I compared the relation between “alive” and “dead” with that between “married” and “divorced.” “X is dead” entails “X has died,” and therefore “X used to be alive,” precisely as “X is divorced” implies “X has been divorced” and therefore “X used to be married.” This is the nerve of the argument. It evokes an endless regress of lives and deaths, in perpetual alternation, which fits exactly the interpretation of “immortal” ( $\alpha\theta\alpha\nu\alpha\tau\omega\varsigma$ ) defended in section I above. Of course such a regress is paradoxical. But the way to sterilize the paradox is to insist that “life” and “death” are not “opposites” in Barnes’s stipulated exhaustive sense. “Dead” and “alive” are no more exhaustive than are “married” and “divorced.”<sup>14</sup> “Dead” does not mean “not alive” or “in-

animate" any more than "divorced" means unmarried." Hence, the claim that the living come from the dead is as implausible as the claim that the married come from the divorced. It follows that Barnes's definition of "opposites" will not fit "life" and "death" any better than several of Plato's other examples.

Suppose, therefore, that we try a different definition of "opposites." Let us count as "opposites" properties that cannot both be true of a subject simultaneously, but may both be false. With this definition, there is no difficulty about counting "beautiful" and "ugly," "large" and "small," "hot" and "cold," or "alive" and "dead" as opposites. But the Principle of Opposites itself now becomes dubious. For, as many have pointed out, it is simply not true that opposites in this sense come from one another. As Barnes nicely observes: "Not all just men rise from an initial state of depravity; not all beautiful swans were once ugly ducklings. I may become just, never having been unjust; I may become ugly, never having been beautiful" (405). So here we have an interpretation of "opposites" which better accommodates Plato's examples, but which unfortunately falsifies his supposed Law.

What is crucial, then, is how we define "opposites." If we take them, with Barnes, as properties that are mutually exclusive and jointly exhaustive then the principle that whatever one acquires one must previously have possessed the other seems correct; but it is not, in the usual modern sense, a "law" and our ordinary concepts of life and death do not fall within its scope. If, on the other hand, we define opposites as mutually exclusive, but not jointly exhaustive, so that life and death fall within their scope, then the Principle of Opposites becomes a lawlike statement, but one which is simply false. In short, Socrates cannot have it both ways. Either opposites do come from each other, but then life and death are not opposites. Or life and death are opposites, but the supposed "law" does not hold true either in general or in this particular case.

(2) Equally important is the Principle of Existence. Barnes gives this (407) as "If anything that exists at a time  $t$  acquires a property at  $t$ , then it existed during a period immediately prior to  $t$ ." He then applies this Principle to the case of the soul's acquiring the property of being alive at time  $t$ . It follows from the Principle, combined with the Principle of Opposites and with the assumption that Life and Death are opposites, that the soul exists and possesses the property of being dead immediately prior to  $t$ . Hence "the souls of men exist in Hades" (408-409). Unlike the Principle of Opposites, the Principle of Existence is nowhere explicitly stated in the text, but has to be extrapolated from the particular case it is

designed to cover. It is not established as a general principle by adducing examples. Rather, it is held to be implicit in Socrates' lines at 70c8-d2, construed as "a special case of some general proposition about change," a proposition which Plato might be imagined to have taken as a self-evident truth (406).

These lines will bear re-examination. I translated them: "if living people are born again from those who have died, surely our souls would have to exist in that world? Because they could hardly be born again, if they didn't exist." Barnes objects to translating  $\gamma\iota\gamma\nu\epsilon\sigma\theta\alpha\iota$  as "be born." He himself translates: "if living people *come into being* again from those who have died, surely our souls would have to exist in that world?" Because they could hardly *come into being again* if they did not exist" (406, my italics). Elsewhere, however, (410, n. 30) he says that  $\gamma\iota\gamma\nu\epsilon\sigma\theta\alpha\iota$  should always be translated "come to be." It will be worth exploring each of these alternatives.

Barnes's actual argument makes "come to be" preferable. For it requires us to read  $\dot{\epsilon}\gamma\iota\gamma\nu\omega\tau\omega$  at 70c8-d2 as incomplete, and to supply "alive" as its complement. Thus the meaning is that our souls could not *come to be alive*, i.e., acquire the property of life, unless they had already existed before doing so. It is from the sentence as thus understood that Barnes extrapolates the Principle of Existence.

It would, I think, be implausible to supply the predicate "alive" for  $\gamma\iota\gamma\nu\epsilon\sigma\theta\alpha\iota$  in many other places where it has no special relevance. A nearby example is 73a2, "unless our souls existed somewhere before  $\gamma\iota\gamma\nu\epsilon\sigma\theta\alpha\iota$  in this human form," where, even if we translate  $\gamma\iota\gamma\nu\epsilon\sigma\theta\alpha\iota$  noncommittally as "coming to be," there seems no reason to understand "alive" as its complement.

At 70c8-d2, where the predicate "alive" is relevant, Barnes's version might seem more defensible. It does not, however, fit at all well with the conception of soul generally at work in the dialogue. For the soul is elsewhere conceived not as *acquiring* the property of life that it previously lacked, but rather as *bringing* to the body the life that it always possesses.<sup>15</sup> It "comes to be present" in the body, bringing life to it. Thus Socrates will ask later (105e9-10) what it is that, when it comes to be ( $\dot{\epsilon}\gamma\iota\gamma\nu\eta\tau\alpha\iota$ ) in the body, the body will be alive. Here, clearly, a complement for  $\dot{\epsilon}\gamma\iota\gamma\nu\eta\tau\alpha\iota$  is neither necessary nor possible. As in the parallel uses of the same verb just above (105b9, c3, c5), the construction is absolute, and we could not understand it to mean "come to be alive in the body." Rather, the soul is thought of as already alive before it enters the body: it "comes to" the body "bringing" life with it (105d3-4), and its

capacity to impart life proves that it will never admit life's opposite. If this is correct, Socrates' words at 70c8-d2 may be taken, similarly, to mean that the soul "enters" or "is born into" the body. They will not then warrant Barnes's "Principle of Existence." They neither mean nor imply that if a thing acquires a property at time  $t$  then it must have existed immediately prior to  $t$ . For they neither say nor imply anything about a subject's "acquiring a property" at all.

Barnes's other version of the key words is also hard to defend. It runs: "They could hardly *come into being* again if they did not exist." But how could we make sense of this, if it means "they could hardly come into existence again, if they did not exist"? Coming into existence would more readily be thought to imply previous *nonexistence* than previous existence. Given Plato's Parmenidean background, an inference from a thing's coming into existence to its immediately prior *nonexistence* would be plausible, whereas an inference to its immediately prior existence would make no sense.<sup>16</sup>

I conclude that both variants for the translation I gave are unsatisfactory. Hackforth (59, n. 3) was right to note at 70c8-d2 that here, as often in the sequel,  $\gamma\acute{e}\gamma\vee\sigma\vartheta\alpha\iota$  "is used not of an absolute coming-to-be of the soul, but of its 'birth' in the sense of incarnation." "Be born" is indeed a familiar and frequent meaning of the verb, as is "birth" for the noun  $\gamma\acute{e}\nu\sigma\tau\varsigma$ . At 88a Cebes will say, with reference to reincarnation, that "nothing prevents the souls of some . . . from being born and dying over and over again—because soul is so strong that it can endure being born repeatedly." It therefore seems natural to connect the verb with the idea of "being born into the body." By understanding it in this way at 70c8-d2, we make the concept as used here consistent with its use elsewhere, and we make the Cyclical Argument conform to the reincarnation mythology that it exploits. But we do not secure Barnes's Principle of Existence.

Unfortunately, to assume the possibility of incarnation is to take for granted what the Argument is supposed to be proving. It is precisely at this point that it seemed to me fundamentally fallacious, and thus here that I accused Plato of "begging the question." I said (105) that the use of "the soul" as subject of "be born" is logically suspect: "For it insinuates a view of 'birth' in which the soul's discarnate existence is already covertly assumed. And since this is precisely what the argument purports to prove, the very concept of incarnation can be seen to beg the essential question." This still seems to me correct. Barnes's interpretation of the Argument allows Plato to escape that charge. But it does so only by sup-

plying him with a general principle that I can find neither stated nor implied in the text.

I asked (109) whether there is such a process as "coming to be alive," and I suggested that in a sense living things may be said to "come to be alive" at birth or conception, but that these events coincide with their coming into existence, and therefore it cannot be inferred from their coming to be alive at birth or conception that they previously existed. Birth or conception, I suggested, is the beginning of their existence (cf. also 171).

Barnes takes me to task for some looseness with regard to "coming to be alive," "birth," "conception," and "coming into existence." He rightly objects that these terms all have different meanings, observing that conception, quickening, and birth occur at different times, and that the Society for the Protection of Unborn Children does not hope to defend nonexistent people from misfortune" (410, n. 30).<sup>17</sup> But it might still be argued that things come to be alive and come into existence precisely when they are conceived;<sup>18</sup> and consequently that no inference to prior existence before coming to be alive or being conceived would be valid. Prior to their conception, living things do not exist. To avoid this result by postulating a preexistent subject, namely "the soul," to acquire the attribute of life, still appears to be question-begging.

To summarize. The Principle of Existence is neither explicit nor implicit in Plato's text. If we accept it, then the discarnate soul must be postulated to serve as the preexistent subject which acquires life. But this postulate is question-begging, and is at odds with the idea of the soul's "bringing" life to the body that it occupies. If we do not accept the Principle of Existence, then we may reasonably believe that we acquire life precisely when we come into existence, at the time of conception. But in that case, no inference from our acquiring life to the prior existence of "us" or "our souls" would be warranted.

#### IV

It remains to consider the Cyclical Argument in relation to Plato's larger design. I suggested (104) that it should be viewed "as an opening dialectical move rather than as an argument to which Plato was seriously committed." Barnes, however, takes the argument seriously, maintaining (401, n. 1) that it is "self-contained" and disputing my suggestion (103) that the argument is refined or superseded later in the dialogue. He holds

that "Plato intended us to take the Argument seriously; and that we have as much reason for supposing that Plato himself was seriously committed to it as for supposing that he was seriously committed to any of the other arguments in the *Phaedo*" (399).

We might distinguish between "taking an argument seriously" and "being seriously committed to it." I did not mean to suggest that the Cyclical Argument should be dismissed as trivial. By calling it "an opening dialectical move," I meant simply that although it purports to prove the immortality of the soul, its weakness will emerge in the ensuing discussion, so that it needs to be repaired or replaced by a better argument in due course.

At 70d4-5 Socrates says that if it is not true that the living come from the dead and from nowhere else, "some other argument would be needed." "Some other argument" might mean a further argument additional to the one offered here, or it might mean, as at 106d1, a different argument to replace this one. Socrates' language and the wider context suggest that he means the latter: the Cyclical Argument would be worthless, and would need to be abandoned in favor of a new one, if it were not the case that the living come only from the dead. But since that premise is established in what follows, and is nowhere withdrawn, it might seem that there is no question of scrapping the Cyclical Argument in favor of a different one. Even if supplemented by additional arguments, as it is, it can still stand on its own feet.

Nevertheless, two considerations tell against the adequacy of the Cyclical Argument "in Plato's mind." First, the Argument strongly suggests an endless alternation of souls between an incarnate and a discarnate state.<sup>19</sup> This suggestion is not, however, strictly consistent with the idea found later, that some souls will escape the cycle of death and rebirth altogether.<sup>20</sup> At the end of the closing myth it is said that "those who have been adequately purified by philosophy live bodiless for the whole of time to come" (114c4), picking up an earlier affirmation that the purified soul "passes the rest of time in very truth with gods" (81a9). Socrates himself speaks as one who expects to enter the company of gods (63b-c, 80d, cf. 115d3-4), as if he exemplified the soul purified by philosophy, who will be with gods for all time, rather than a soul that must wait "in Hades" until it is born again. "Hades in the true sense" (80d), which he hopes to enter, contrasts, as we have seen, with the Hades of myth. If Plato wished to invert the traditional concept of Hades, should we give as much credence to the Cyclical Argument which draws upon it as to Socrates' personal expectations?

Secondly, serious doubts about the worth of the Argument are implied later, notably at 77b1-c5, where Simmias and Cebes both complain that our souls' postmortem existence remains unproved. Even though Socrates purports to deal with their difficulty, he rightly suspects (77d5-e2) that they will remain unconvinced. Both of them subsequently press objections that would, if sound, seem fatal to the postmortem part of the Cyclical Argument. At 91c they claim to reject "some" of the previous arguments but not others. Since only three arguments altogether have been given, and one of them ("Recollection") has convinced them both, it surely follows that they reject both the others, and not just the Affinity Argument to which their objections were directly addressed.

Barnes replies that these doubts "are not shared by Socrates, and we have no reason to suppose that they were Plato's own doubts" (399). But at 84c, even before Cebes and Simmias have urged their objections, Socrates himself says: "what's been said certainly still leaves room for many misgivings and objections, if, that is, one is going to examine it adequately," and similar hesitations are voiced elsewhere (85c-d, cf. 86d7-8, 107a-b). These remarks, admittedly, do not refer specifically to the Cyclical Argument, nor is that Argument ever directly attacked. But the objection of Cebes, in particular, suggests precisely where the Argument was defective in Plato's view. Let me enlarge on this.

First, whose immortality is being proved? The answer is clearly "our souls'." I agree with Barnes that what is at issue throughout the Cyclical Argument, as in the whole dialogue, is the fate of the individual soul, rather than any kind of cosmic or depersonalized soul. Barnes is thus right to reject the conclusion of Friedländer (400, n. 10) that "it is not a matter of my death or yours, or of your life or mine."<sup>21</sup> But even so, we might wonder whether "our souls" means those of the present speakers, or whether "our" and "we" in this context simply refer to the souls of the human race generally.<sup>22</sup>

It is reasonably clear that when Socrates says, during the Cyclical Argument, "our souls are in Hades," he means the souls of human beings quite generally (cf. "the souls of men," 70c5). He does not at this point refer specifically to the souls of present company at the present time, since their souls are obviously not in Hades at that time, and a present tensed "are" with reference to them would be inappropriate.

Much later on, however, Plato makes Socrates use the future tense, "our souls *will* exist in Hades" (107a1). This is a significant shift, for it permits "our souls" to refer here with special relevance to those of the present speakers. They can be assured, in the light of the foregoing ar-

gument, that their own souls *will* survive death. For a major objection pressed by Cebes (88a-b), or by Socrates on his behalf (95c-d), can now be represented as fully met. It was his concern that even if a whole series of incarnations were granted, that would give the individual no proof of the survival of his own soul. So the future tense enables the phrase "our souls" to refer to those of the present company, as befits the dramatic situation.

Now what has the Cyclical Argument proved with respect to the future? There is an ambiguity about the concept of "the future" to be noticed here, which also affects such phrases as "after death," "when we have died," "postmortem," and the like. If one wishes to argue with respect to a whole class of Xs what they will do "in the future," one needs to distinguish the future of previous Xs from those of present or subsequent ones.<sup>23</sup> In claiming, for example, that caterpillars will turn into butterflies, one may mention previous caterpillars that have done so already. Of them it would once have been correct, when they were caterpillars, to say that they would turn into butterflies. Similarly, Socrates might prove the "future" or "postmortem" existence of the soul, by showing that some souls must have survived death in the past. Of them it could at one time have been truly said that they would exist "in the future" or "after death." In that limited sense the Cyclical Argument proves the "future" or "postmortem" existence of any souls that are animating living things at whatever time the Argument is being deployed. The proof works in retrospect: they must have survived their previous incarnations, in order to be reborn in their present ones.

Such an argument, however, would prove the future existence of all souls, and therefore of our own individual souls, only if an inductive inference from the futures of previously incarnate souls to the futures of presently and subsequently incarnate ones were sound. But it is precisely the soundness of this induction that Cebes calls in question. How, indeed, can one move from the future of previously incarnate souls to the future of present or subsequent ones? If "the future" is construed from the perspective of speakers in this discussion, rather than from that of souls in a previous incarnation, it is doubtful whether the Cyclical Argument has proved anything about "the future" at all. At this point, the supplementary argument at 77d3-4, that the soul "must also exist after it has died, given that it has to be born again," is extremely feeble, and at least in the case of a soul like that of Socrates its premise would be false. Cebes' objection fastens unerringly upon this crucial weakness. And since Plato is the author of Cebes' words, as well as of the Cyclical Argument, he

must surely have perceived the weakness himself. It is, presumably, to remedy it that he will make Socrates deploy his "Final Argument," to prove the soul "completely immortal and imperishable" (88b, 95b, 106e-107a).

How, then, should we determine the "seriousness" of the Cyclical Argument? Perhaps as follows: For all its defects, Plato might have thought that the Argument proves both the prenatal existence of souls now or previously incarnate, and the postmortem existence of at least some souls previously incarnate, viz. those that are incarnate now. But he did not think it proved the postmortem existence of every soul in every incarnation, hence of all souls that are now incarnate. On that point, at least, he needed "some other argument," and it is the Final Argument that provides it. Either it repairs the Cyclical Argument by supplying a proof of postmortem existence at the point where the Argument fails; or else it supersedes the Cyclical Argument altogether, by rendering it unnecessary. On either view it possesses a finality and seriousness that the Cyclical Argument lacks.

One last point should, however, be stressed, familiar though it is. In a certain sense none of the dialogue's arguments is "final" and none is to be taken completely "seriously." All of them are exploratory, and, like the written word in general, are composed "in play" rather than "in earnest" (cf. *Phaedrus* 276b-e). This applies to the *Phaedo* as a whole. Barnes characterizes it as "a philosophical treatise of power, originality, and influence" (397). With the last five words no one could quarrel. But "treatise" seems to me the wrong word for the dramatic masterpieces of Plato's maturity. For they are not expositions of doctrines so much as explorations of philosophical issues. In pursuit of those issues, the Cyclical Argument, whatever its defects, remains a powerful stimulus. In that respect it is in no way inferior to the later arguments, but is simply continuous with everything else that follows it.

## Notes

1. Drafts of this paper were read to the Classical Philosophy Seminar at the University of Texas at Austin in April 1981, and in the Ontario Classical Philosophy Workshop sponsored by Trent University in February 1982. I am grateful to members of both groups for helpful discussion.

2. *Canadian Journal of Philosophy*, VIII. (1978): 397-419. References to this critical notice and to my *Phaedo* will be given by page number only.

3. "The Last Argument of Plato's *Phaedo* II," *Classical Quarterly* (N. S.) 18 (1968): 95-106, especially 96-97; see also R. Hackforth, ed., *Plato's Phaedo* (Cambridge: Cambridge University Press, 1972), 163-65.

4. For further examples, see L.S.J. s.vv. *ἀθάνατος*, *ἀθανασία*.

5. Aristotle's usage is also of interest. When he instances *ἀθανασία* as an object of wish for the impossible (*N.E.* 1111b22-23, cf *E.E.* 1225b33-34), he means permanent immunity from death, rather than survival of the disembodied soul. At *N.E.* 1177b33 *ἀθανατίζειν* means not survival of death but "playing the immortal" during one's lifetime. At *Topics* 145b21-23 *τὸ ἀθάνατον* is equated with "being such as never to be destroyed." This passage is cited by O'Brien ("The Last Argument of Plato's *Phaedo* II," 97, n. 2), but as he recognizes, it identifies *ἀθάνατος* with *ἀφθαρτος* in the strongest possible sense. It therefore lends no support to the idea of a "partial" form of immortality.

6. Secondary forms of immortality are, of course, recognized by Plato, notably the vicarious *ἀθανασία* achieved through the generation of children or other kinds of offspring (*Symposium* 207a-208b). The comparative form *ἀθανατώτερος* occurs twice (*Symposium* 209c7, *Phaedo* 99c4), but neither text evidences "partial" immortality in the sense discussed above.

7. See my note on 96b5-c6 (168).

8. Nor is there any suggestion, either here or elsewhere, that the soul might perish at any time other than when it separates from the body. As O'Brien points out, "Plato consistently supposes that if the soul were to become extinct it would do so at the time of its separation from the body. This is the only time at which the soul is thought to be vulnerable" ("The Last Argument of Plato's *Phaedo* II," 100 with nn. 1-2).

9. See my note on 77a6-d5 (135-36).

10. Cf. 76c11-13, where (A) is restated with respect to prenatal existence only. Barnes's view that the Cyclical Argument is restricted to prenatal existence, taken with the fact that the Recollection Argument must be thus limited, yields the following strange result. When Simmius and Cebes object at 77b-c that the soul's postmortem existence has not been proved, Socrates replies that it *has been proved already* if the two arguments are combined: somehow, then, two arguments which, on Barnes's view, purport to prove only prenatal existence if taken separately, are alleged by Socrates to have proved *post mortem* existence already if taken together!

11. Orpheus is recognized as an exception at *Symposium* 179d6-7. It may be of interest to note one difference between the underworld of *Odyssey* XI and that of *Aeneid* VI: the latter (752 ff.) contains the shades not only of those who have already lived but of Roman heroes yet to be born. This was presumably dictated by the patriotic purpose of Vergil's epic, and has no parallel, as far as I know, in Greek conceptions of Hades.

12. Teiresias is a partial exception, mentioned as such at *Meno* 100a. We should recall that the relevant passages from the *Odyssey*, and similar texts from the *Iliad*, are to be withheld from the young guardians in the *Republic* (386a-387b). This is consistent with Plato's wish to stand the traditional picture of Hades on its head.

13. See my note on 70c8-d6 (107).

14. My colleague Trudy Govier has pointed out to me that "married" and "divorced" are not exclusive either. For, unlike "dead" and "alive," they are relational: X may be married to A and divorced from B.

15. Barnes says (413) that the theory suggested in the *Phaedo* is that "I (or my soul) become alive at certain times, and becoming alive consists in coming to animate a natural body." This does not seem to me to express the dialogue's dominant view. The soul is alive at all times, and not just when it animates a body. There is, in fact, a conscious extension of the concept of "life" beyond the confines of "what we call life" (107c3).

16. Weight might perhaps be placed upon the word "again." The argument might then be that our souls could hardly come into existence *again*, unless they had existed at one time already (and then presumably ceased to exist). But it is doubtful whether the words will bear such weight. Moreover, on that view we would have to posit a period of nonexistence for the soul, intervening between its supposed earlier existence and its coming into existence "again." But the idea of such an intermission is quite alien to the view of the soul developed in the *Phaedo*: there is no question of its nonexistence, or of its coming into existence at *any* time. Cf. *Phaedrus* 246a1, where it is characterised as *ἀγένητον* as well as *ἀθάνατον*.

17. It might be added that Plato could distinguish, where he wished to, between "being conceived" and "being born." Cf. *Symposium* 203b-d, *Euthydemus* 296d1. Moreover, in some places *γίγνεσθαι* could not possibly mean "be conceived," e.g., 75b10, where it is said that we saw and heard and possessed our other senses "just as soon as we were born" (*γενόμενοι εύθυγάτοι*).

18. Alternatively, it might simply be denied that there is such a *γένεσις* as "coming to be alive" or "acquiring life" in the normal course of events. Neither in Greek nor in English is "coming to be alive" synonymous with "being born" or "being conceived." The phrase "comes to life" finds a natural use only in those abnormal cases where the subject is an inanimate and observable item, such as a corpse or a statue. There are no problems of reference in saying of such items that they "come to life"; but there are other grounds for doubting whether they could in fact do so, and consequently for doubting whether there is such a *γένεσις* as "coming to be alive." To raise this doubt is to emphasize once again how critical for the Cyclical Argument is the principle stated at 71a12-b2 that *γένεσις* between every pair of opposites occur in both directions. See my notes on 71a12-b5 and 72a11-d5. Barnes has an interesting alternative explana-

tion (405-406) for this part of the Cyclical Argument, which I do not, however, find convincing.

19. Exactly this idea is found also in the *Meno* (81b3-6).

29. Cf. also *Phaedrus* 248c5.

21. See my note on 70c8-d6 (106).

22. The reference of "we" varies in the *Phaedo* between all of those present (88c1-7), the speakers in the discussion (103a5), the wider community of philosophers (75d2), the users of the Greek language (107c3), and the present speaker (116d4, and perhaps 69d5-6, 118a7).

23. F. L. Will has drawn a similar and relevant distinction between senses of "future" in "Will the Future be like the Past?" *Mind* (56) 1947, 332-47, reprinted in *Logic and Language, First and Second Series*, ed. A. Flew (Garden City, New York: Doubleday Anchor, 1965), 248-66. See especially 262-66 of the reprinted edition.

## 12

### The Final Proof of the Immortality of the Soul in Plato's *Phaedo* 102a-107a

Dorothea Frede

Among the arguments presented by Socrates as proofs for the everlastingness of the human soul the last one has greatly puzzled philosophers because it seems that, in opposition to the earlier arguments,<sup>1</sup> Plato considered this last argument conclusive. For, whereas earlier in the discussion the partners of the dialogue, Simmias and Cebes, raise objections and Socrates tries to meet their criticism, at the end of the last argument he claims: "Then this is most certain that the soul is immortal and imperishable and that our souls will really exist in Hades" (106c). And at this point Socrates obtains his partners' final consent. For, though Simmias admits that he still entertains doubts, it seems that Plato wants to attribute this to the natural difficulties we encounter when *arguments* have to overcome *fears*, not to shortcomings of the argument itself (cf. 107a-b). For with Socrates' calm reassurance that further scrutiny of its presuppositions would finally lead to a state of conviction as far as this is humanly possible, the argumentative part of the dialogue ends (107b9).<sup>2</sup> What follows is Socrates' mythical description of the soul's afterlife and, finally, the narrative of Socrates' peaceful death among his friends.

Because of Socrates'/Plato's apparent confidence in this last argument it has always attracted special attention. And in recent years there have been various attempts to expose the flaw or flaws in the proof. For, ever since Kant pointed out the unavoidable difficulties that human reason encounters when it tries to transcend the boundaries of sense-experience, the interest of philosophers in arguments of this kind has been mainly a critical one. It is a challenge to a philosopher's skill to ex-

pose weaknesses of such arguments, steps which do not follow necessarily or assumptions which one does not have to accept. Furthermore, especially in the case of a philosopher as great as Plato, such a critical investigation promises to be a rewarding task, since one may hope that, if the argument under investigation turns out not to be clearly invalid, one may learn something about those philosophical presuppositions which are not explicitly stated in the argument but tacitly assumed (or argued for earlier in the dialogue), but which are, nevertheless, needed to make the argument go through. In this paper, then, I shall try to defend Plato's argument against those critics who claim that its conclusion is reached by invalid inference and point out which assumptions Plato may have taken for granted.

The final argument is introduced after what may be considered as a long interlude dealing with the problem of the notion of cause or reason (cf. 95e-99d). This interlude had been prompted by Socrates' admission that in order to meet Cebes' criticism he would have to go into an explanation of the causes of generation and destruction (95e-96a). This, however, Socrates felt himself unable to do because the notion of cause or reason presented such difficulties to him that he finally took refuge in what he calls a "δεύτερος πλοῦς," a second-best way, i.e., the resort to the forms which at least allows him to explain why the things are the way they are.<sup>3</sup>

That Socrates is turning back to the question of the immortality of the soul after this interlude is not immediately obvious, for what follows seems at first sight to be merely a restriction and clarification of the so-called law of opposites (or—"alternation," a principle Socrates had used in his first argument for the immortality of the soul 70c-72a: Wherever there are opposite conditions and processes through which things pass from one to the other, there must always be exchange in both directions). This is here clarified and modified (102b ff): not only do things pass from one state into its contrary but some things can even be in opposite states at the same time; as, e.g., Simmias is at the same time both taller and shorter—taller than Socrates and shorter than Phaedo.<sup>4</sup> Yet, although Simmias himself can participate in opposite forms at the same time, neither the form "tallness" itself nor the immanent forms or characters can do the same; they are incompatible with each other and therefore, as Plato expresses himself throughout the argument in terms of the military metaphor, they "either perish at the approach of their opposite or they withdraw" (102d-e; 103a1 et pass.; cf. Burnet's comment *ad loc.*). So, although Simmias himself can become taller or shorter while still re-

main Simmias, his tallness or shortness cannot accept their respective opposites.

The bewilderment expressed at this point (103a) by someone in the audience shows why Plato thought this clarification of the law of opposites necessary. There is an easy misunderstanding, especially since the Greek "τὸ μέγα" can be understood both as "tallness" and as "that which is tall"; to the ordinary listener it might have sounded as if the opposite states were supposed to turn into each other, while it is here explained that it is only something *in* such a state which can pass from one to the other (103b-c).<sup>5</sup>

What follows is a further reconstruction of the law of alternation (103 ff). In *some* cases it is not only the opposite states or qualities and their forms that cannot turn into their opposites but even their possessors—though not identical in nature with those characters—cannot accept the opposite of the characters they possess; they either have to withdraw or to perish, too. Snow, while being snow, never admits heat; and fire while being in existence never admits cold. So, though these two things, snow and fire, are not identical with the properties in question, they still cannot adopt the qualities opposite to those they do in fact possess. Plato, as we would say it, distinguishes here between the possession of essential and accidental properties and modifies the law of opposites accordingly.<sup>6</sup>

The distinction between essential and accidental qualities allows Socrates to proceed one step further. Not only does the possessor of an essential quality not admit its opposite, but he can even be *called* after that incapability of admittance: "three" is incapable of admitting evenness and can therefore be called "uneven" (104e5). So, whenever something has an essential property it can be called unqualifiedly after that property (and negatively after its opposite). And, in such a case, whenever someone wants to know *why* a thing is P one has only to point out that it is possessed by something which has P as its essential property—"If you were to ask me what, when present in its body will always make a thing hot, I shall not give you that safe, foolish answer, 'heat,'" but, after what we have just said, a more subtle answer, 'fire' . . . and if you ask what, when present in a number will always make it odd, I shall not answer 'oddness' but 'oneness' and so on" (105b-c).

The distinction between accidental and essential properties and the acceptance of the possibility that such properties can be "brought along" by their possessors to some third thing then leads over to the final steps in the argument for the immortality of the soul, which is now brought to

its end in a very brief discussion (105c-d-107 a).

There are, roughly speaking, three steps in the final discussion:

1. Socrates obtains the agreement of his partners that the thing which always makes the body alive when it is present in it is the soul. Thus, soul always brings life with it when it approaches something (105c9-d5).

2. Since life is the opposite of death, the soul cannot admit death. And, in analogy to the other cases where there existed such an unqualified incapacity to accept a quality, we can call the soul after its incapacity to admit death "deathless" (105d6-c9).

3. If in the analogous cases the specific incapacity to accept something also implied the incapacity to accept destruction, then those things could never be destroyed. They would be unassailable to attack by their opposite principle (snow by heat, fire by coldness, three by evenness) but would withdraw intact. This, however, does not hold for the analogous cases. But since for the "deathless" it is accepted that it is at the same time also indestructible, it follows that at the approach of death only what is mortal in man dies; the soul never perishes but withdraws safely to Hades (105d6-107a1).

These, without being unfair to Plato's more subtle and elaborate way of presenting his argument, seem to be the main steps, and all three of them have been attacked by Plato's critics. Let us, first, look ourselves at possible weak points before turning to other criticisms.

It may seem that Plato's claim that the soul always carries life with it is simply begging the question. For this very point seems to be what is at issue in the whole argument: whether the soul *always* carries life with it; and it is here not even argued for but simply granted at Socrates' suggestion by his partner (105d1). What can be the justification for this assumption? —A justification is suggested by the fact that what is emphasized in step I is that the soul always quickens the *body* when it is there, and that therefore its possession of life cannot be an accidental one. For in that case it should at least be logically possible to say that there still is a soul in a certain body but that nevertheless the body is dead. This, however, seems an impossible assumption, just as it is impossible to maintain that something is possessed by fire but not hot or that something is a threesome but not odd.<sup>7</sup> So the first step does not rest on the assumption that dead souls are inconceivable,<sup>8</sup> but on the general rule given in 104d1-3 for finding entities possessing something as an essential property: whenever something always imparts something else, it cannot

possess this thing only as an accidental property. For if it were an accidental one it should at least be conceivable that it might not bring along the property in question. Thus, the first step in the argument has to be accepted if one accepts the general rule in 104d (certain difficulties about the text and the application of the rule will be discussed later), i.e., one has to accept that the soul always possesses life while in existence.

Let us, then, look at step II. If the soul has life as its essential attribute, then it can be called, after its incapacity to accept the opposite, "deathless." To be sure, the word " $\alpha\vartheta\alpha\nu\alpha\tau\omega\varsigma$ " is as ambiguous as the English word "immortal," which has been carefully avoided by the English translators since it designates not only deathlessness but also everlastingness. But it does not seem that Socrates is trying to win an easy game by way of this ambiguity here, even though his partners might have been willing enough to let him win it (cf. 105e). Instead, he makes it quite clear that the validity of the argument depends on the condition that the deathless should also be indestructible (105e10 ff), i.e., that the soul cannot be destroyed in any other way. Therefore, if we agree that the soul can be called "deathless" after its essential incapacity—just as other things are called after theirs—step II is legitimate and we would have to look at step III for possible criticism, at the thesis that the deathless is also indestructible, and that therefore the soul must be everlasting.

It is, unfortunately, not easy to elicit from the text (106b-e) how step III is established. Many commentators, starting as early as Strato of Lampsacus, have accused Plato of begging the question here, i.e., of simply assuming what needs to be proved, that the soul, being "deathless," cannot go out of existence in some other way. Is Plato simply assuming that deathlessness = indestructibility = everlastingness?

There are at least some indications in the text that Plato did not commit this obvious mistake. What speaks against it is the fact that Socrates repeats three times that the validity of the last argument depends on the question whether “deathless” also implies “indestructible” (cf. 106b2 “εἰ μὲν τὸ ἀθάνατον καὶ ἀνώλεθρόν ἔστιν”; 106c8 “περὶ τοῦ ἀθανάτου, εἰ μὲν ἡμῖν ὅμολογεῖται καὶ ἀνώλεθρον εἶναι . . . εἰ δὲ μὴ.”) And even after his partner has given his final consent Socrates repeats: “ὅπότε δὴ τὸ ἀθάνατον καὶ ἀδιάφθορόν ἔστιν” 106e1). Plato would hardly have repeated this important condition three times if he had simply taken it for granted that deathlessness also means indestructibility. And earlier Plato had stressed that all the entities only possess their essential attributes “ὅταν περ ἦ” while they exist (103e),<sup>9</sup> a condition which is again stressed in the text for all the analogous cases

in step III, that they would only survive the attack of the opposite of their essential attributes if they were also indestructible.

Something of a clue as to Plato's reasons for inferring indestructibility from deathlessness is given when in 106d Socrates not simply accepts Cebes' assent that the deathless is indestructible but adds a reference to "God, the form of life itself, and if there is anything else immortal": that it would be admitted by everybody that they could never perish. If this is not a rather pointless appeal, there must be something which God, the form of life and the unnamed "etceteras" have in common (especially in the case of the "καὶ εἰ τι ἀλλό" in 106d6, a mere commonsense appeal would seem strange). This common characteristic must be that they are all *essentially* alive.<sup>10</sup> And in this case the law of opposites will find its legitimate application. This law had never been explicitly challenged in the dialogue (although this does not exclude the possibility that Plato had been aware of its weaknesses as a general law). It is explicitly referred to and its meaning and applicability are clarified at the beginning of our last argument (103a4-c4), so it would not be excessive subtlety on Plato's part if he supposed that his readers would realize that it finds its correct application here.<sup>11</sup> For, whatever is alive, whether it possesses life as an essential property or not, can only pass out of existence by accepting death, by dying. And, whatever way a living being passes out of existence we call that its death, be it a "natural" one or, e.g., traceless elimination through an atomic bomb. So it is quite inconceivable that to the question "When did N. N. die?" one would receive the answer, "He did not die, he simply went out of existence."

It is easy to point out why the law of opposites does not hold as a general law. It is valid in the case of relative opposites, like "greater" and "smaller," the kind of cases from which Plato had set out (70e7-71a7), apparently without noticing that and why these are special cases. In the case of absolute opposites the law holds in some cases but not in others, as becomes clear when one looks at some examples. While someone who has fallen ill must have been healthy before and vice versa, not everybody who has become rich must have been poor (or vice versa). In some cases we will find that while one half of the process is necessary the other is not. If someone becomes old he must have been young before, but the young people do not come from the old ones. That is to say, the process may be reversible, as in the case of waking and sleeping, but it need not be so; as Plato seems to assume, relying heavily on the poetically much exploited analogy between waking and sleeping and life and death (71c-d). For, even if creation *ex nihilo* is excluded there are still

several other possible explanations of how people come to life, and hence there is no necessity to assume that they come from the dead. Soul may come, e.g., from other souls, just as fire comes from fire. This criticism of the law of opposites as a general law, however, does not apply to the usage Plato makes of it in the final argument. For in the case of life and death the reverse side is necessary, in other words, whenever something loses its life it *must* pass into death. And Plato seems to me to refer to this fact when he points out, as he does repeatedly, that the validity of his argument depends on the condition that that which cannot die cannot be destroyed. For if destruction for a living being is its loss of life (death), then deathlessness implies indestructibility. Thus the inference is justified that whatever possesses life as essential attribute cannot be destructible; for if it cannot admit death it cannot go out of existence at all, and must therefore be indestructible as well.<sup>12</sup>

So, if I am not mistaken in my reading, Plato's argument for the immortality of the soul is *formally* correct, i.e., Plato has neither drawn false inferences nor simply begged the question. People who nevertheless are not convinced that the soul is immortal or that its immortality can be proved therefore will have to attack the premises of the argument. Which of the premises is it at least not necessary for everybody to accept?

The critical point should be obvious, since in my interpretation of the argument I had to use a fairly suggestive vocabulary. The crucial thing seems to be that Plato treats the soul as a substance (no such term is, of course, used) with attributes of its own and life among them. But this is a presupposition about the nature of the soul which one may or may not find acceptable. There are different accounts of the nature of the human soul and Plato's arguments in the *Phaedo* do not rule them out.

In other words, Plato's argument would be acceptable to people who believe with him that a human being is really a compound of two entities, the body which is the material vessel, and in it the incorporeal soul which has essential qualities of its own. The body, then, incidentally shares in the quality "life" just as a stove which is heated by the fire inside it shares in the quality "heat." So Plato's argument is valid as a proof only if everybody had to accept the hypothesis that the soul is an entity like fire, an entity bringing along properties of its own. But if the soul were not a being *with* qualities but a quality itself—the "energy of life"—like the heat in fire or the cold in snow, then soul, the principle of life in the body, could simply run out without "admitting" death itself at all. It is clear that in this case the soul would not be separable from the body nor possess qualities of its own. If the body's functions were so seriously

disturbed that it stopped working, it would just lose the quality of life, and there would be no entity left in this case except the lifeless body.

Admittedly, then, the defense of the formal correctness of Plato's final argument depends on the presupposition that he regards the soul as something like a substance. There are, however, passages in the argument which have given reason to doubt whether for Plato the soul is not rather an immanent form or character bringing along another form, just as threeness brings along oddness. Hackforth (159; 161 ff) and other commentators after him have tried to point out that Plato's position with respect to the nature of the soul is somehow undecided; that during the argument he treats it like an immanent form but concludes that it exists like a separable substance.<sup>13</sup> What speaks in favor of this interpretation is *a) the context of the final argument in the dialogue: in the text immediately preceding our argument Socrates had obtained everybody's approval that the resort to the forms would be the best solution to the question why everything is the way it is, and at the end of the argument Socrates turns again to the forms: further scrutiny of the first hypotheses would lead to final satisfactory conviction (107b); b) the fact that not only does Plato not keep his examples very carefully apart, like that of snow possessing coldness and that of threeness bringing along oddness, but he also illustrates his conception of one thing "bringing along" and "imposing" one of two opposite characters on what it gets hold of (104d1-3) by referring to the way in which the form three takes hold of something, imposing not only threeness but also oddness, on the latter (104d5-e10). It should also be added that, perhaps, for the sake of clarity it would have been preferable if Plato had argued exclusively on the level of the forms and their mutual relations. But it seems clear that he is not doing this.*

There are good reasons, in fact, to suspect that Plato quite deliberately chooses examples of kinds as different as that of fire bringing along heat and threeness bringing along oddness, for they are placed side by side even within one and the same sentence (104e-105a, 105c; and again 106b-c). —That fire and snow are not regarded as forms seems to me obvious. It is true, that in the *Timaeus* Plato does talk about the form of fire (cf. 91b ff)—but only as long as its atomic structure is under discussion, the physical fire is then mentioned along with the other elements as "σώματα" (*Tim.* 53c4; 57c).<sup>14</sup> And of snow it is merely stated that it is a form of water, not a form *in* water. And the way Plato talks in our argument itself about snow immediately suggests that he is talking about the physical entity. For he explains (106a) that if snow in addition to being

essentially cold were also indestructible it would retreat at the attack of something hot "σῶς καὶ ἀτηκτος"—safe and unmelted. But, in whatever way an immanent form of snow may retreat, it would have to leave behind its "corpse" i.e., warm water.<sup>15</sup>

So it seems to me that Plato, who was, after all, completely at liberty to choose his own examples, easily could have avoided such doubtful cases as fire and snow if he had wanted to talk only about immanent forms. That he returns to the examples of fire and snow at all occasions therefore suggests that he thought that no more than an analogy between the various cases was necessary for his argument. Or, to put it like Hartman (218-20), he may have chosen these different kinds of examples quite intentionally in order to show that the applicability of his model did not depend on any specific case. This would also be a good explanation for the fact that throughout the argument Plato sticks to his military metaphors of "attacking," "occupying," or "retreating." For while this can be understood in an almost literal sense in the case of fire or snow, it is not easy to see what its exact meaning is in the case of the immanent characters which may differ from each other as widely as Socrates' size relative to Simmias' and threeness bringing along oddness. This metaphorical way of speaking allows Plato to evade the question of how the immanent characters take over, possess, and leave their occupied entities, and what it means in each case that they "withdraw or perish."<sup>16</sup>

Against this it has been argued that 104d1-3, the passage in which Plato states his general rule for finding entities with essential attributes, in its most natural way of reading would refer to the forms, i.e., that they are what impose their own character along with some opposite on whatever they occupy: ἀ ὅτι ἀν κατάσχη μὴ μόνον ἀναγκάζει τὴν αὐτοῦ ἰδέαν αὐτὸ ἵσχειν, ἀλλὰ καὶ ἐναντίου αὐ τῷ τινός (where I read "αὐ τῷ" with Bluck; cf. 105a3-4). Gallop, who does not want to regard the soul as an immanent form, has therefore tried to defend a different reading of the passage (cf. translation, commentary, and notes *ad loc.*), to the effect that the items Plato talks about in 104d1-3 are the things occupied rather than their occupiers (cf. esp. 203-207). This interpretation, though grammatically possible, has the disadvantage not only of destroying the correspondence between "ὅτι" in line 2 and "αὐτό" in line 3, but one has to accept a certain shift which the text does not seem to warrant: The "items" in question according to G. are at first the things occupied, then those that "bring along" something to something else, and finally those that occupy something else (as in the case of the soul in 105d3-5). Against this, Plato's text suggests that it is the same

kind of entity that “occupies” in 104d1-3, “brings along” 104e-105a, and that the soul is just one of their kind.

In opposition to Gallop I think that one can read the text in 104d1-3 in its most natural way and still maintain that the occupiers (including the soul) are not necessarily forms. For the formulation in 104d1-2 seems to me to be quite neutral as to the kind of entities that are supposed to “κατέχειν” something else: it is merely stated that they impose their own form on the thing occupied without precluding that those entities should be forms or things *with* forms. That the example added is of the form three “ὅτι ἀ ἀν ἡ τῶν τριῶν ἰδέα κατάσχη” (104d5-6) does not decide the matter, since, as said earlier, Plato sees no difficulty in switching back to entities like fire in the same context (105a1).<sup>17</sup>

But, someone might object, if the text is at best neutral on the question, i.e., neither dealing exclusively with forms nor with substances, why, then, should the soul not be a form? That it cannot be an immanent form becomes clear when one reflects on the possible ontological status of immanent forms—they owe their existence to the fact that something participates in the form itself. But how can there be a free-floating immanent form which is neither “ἐν ἡμῖν” nor identical with the form as such “ἐν τῇ φύσει” (as Plato distinguishes them in 103b)? That the soul cannot, on the other hand, be identical with the form as such can be derived from the Affinity Argument, where Socrates had explained that the soul is more *like* and *closer* in kind (80 b) to the intelligible, invariable, and eternal forms (78d) and therefore rather akin to them than to the earthly impure entities, but the soul is clearly not recognized as one of the forms.<sup>18</sup>

Furthermore, as some commentators on the last argument seem to neglect, Socrates can rely for it on certain presuppositions on which he and his partners had agreed as the result of the earlier arguments. Socrates had refuted Simmias' suggestion that the soul might be something like a “harmony”—invisible, incorporeal, divine, but yet depending for its existence on the body (instrument) which possesses it (cf. 85 ff.). Also, both Simmias and Cebes had accepted the argument from recollection—that the soul must have had a prenatal life in which it has had access to the forms (cf. 77a ff.; 92d-e). And all that Cebes demands in his criticism is that the soul may not eventually be worn out after several reincarnations but is really everlasting. And, if the final argument contains Socrates' answer to Cebes, this is the only step that has to be filled in by our argument. Plato may therefore have relied on it that it is agreed that the soul is an independent entity with properties and capacities of its

own, something which we would call a substance.

Before we turn to a further investigation of Plato's conception of the nature of the soul some further criticism or the final argument has to be dealt with. E. Hartman has pointed out that even if the soul is not regarded as a form but as a substance bringing along qualities, the argument fails because Plato wrongly assumes that whenever something “brings along” something else it should possess this as a quality itself. Hartman has no difficulty in showing that this is not true as a general principle. The hemlock brings death to man without being dead itself, fever brings sickness without being sick etc. (221 ff.). This seems, at first, a very serious reproach. For not only does Plato indeed argue that the soul never can take on death because it always brings along life to the body, without giving any justification for passing from “brings along” (ἐπιφέρει) to “does not accept the opposite” (οὐ δέχεται) in 105d-e; but the difficulty seems also to hang together with the assumption which Plato himself later realized to be troublesome for his theory of the forms, i.e., that they all are supposed to possess the character they stand for, in such a way that justice is just and tallness is tall—the problem of self-predication.

Plato has been deluded in our argument, as Hartman explains, by what has been called by S. Peterson “Pauline Predication,” the way in which St. Paul in Corinth. XIII, 4 talks of “charity as being patient, kind and long enduring,” where it is, of course, not charity itself which has all those virtues but the person who possesses charity.<sup>19</sup> Analogously, it seems clear that the soul need not possess life itself just because that which has a soul has life.

Sophisticated and suggestive as this criticism seems, I do not think that it does justice to Plato here. Fortunately, we do not have to rely only on general considerations such as the consideration that an argument for the immortality of the soul makes no sense if it is not at least assumed that the soul is alive before death. I think that, in fact, some of the conditions which Plato introduces in the final argument are quite sufficient to rule out cases of Pauline predication. For the inference against which Hartman directs his criticism, that “that which always brings a character on something else possesses this as an essential characteristic itself” (cf. 104d1; 105d3), is not the only criterion which Plato uses here but only supplementary to a criterion given earlier which rules out “Pauline” characteristics.

First, already the way in which Plato had demarcated the distinction between forms, immanent forms or characters, and the possessors of

those characters in 103b-c suggests that only those cases are under consideration in which the character in question is a genuine “ἐνόν,” and not merely imposed on something else (as in the case of the deathbringing hemlock). Secondly, he adds an even more important condition, namely, that the possessor should be *called* after the character he possesses—“ἐπονομάζοντες αὐτὰ τῇ ἐκείνων ἐπωνυμίᾳ” (103b8-9). That Plato attributed some importance to this “eponymy” criterion is confirmed by the fact that he uses it again when he makes the distinction between accidental and essential attributes: “Ἐστιν ἄρα περὶ ἔνια τῶν τοιούτων, ὥστε μὴ μόνον αὐτὸ τὸ εἶδος ἀξιοῦσθαι τοῦ αὐτοῦ ὀνόματος εἰς τὸν ἀεὶ χρόνον, ἀλλα καὶ ἀλλο τι ὃ ἐστι μὲν οὐκ ἐκεῖνο, ἔχει δὲ τὴν ἐκείνου μορφὴν ἀεὶ, ὅτανπερ οὐ”; (103e; the criterion is again used in the special case of “three” and “odd” in 104a). Now, Hartman’s examples clearly do not fulfill this criterion: neither the death-bringing hemlock nor the man-eating tiger can even accidentally be called after what they bring along to other things.<sup>20</sup>

The further condition “bringing along” is, then, not introduced as a means to find out whether a certain thing possesses a certain character in the first place, but merely as a further criterion to distinguish between accidental and essential possession. For this is how Socrates applies the rule: as a method to determine whether a certain thing *always* possesses a certain property (cf. 104b6 ff; d3).

It is clear why Plato needs this second criterion in the case of the soul. “Eponymy” as such is not sufficient, because, as Plato realized (cf. 103b6 ff; e2), all things can be called after all their properties, even if they are only accidental ones (as long as they possess them in a non-Pauline way); I can be called awake or asleep, healthy or ill, although these are clearly only accidental properties. The addition “can *always* be called” (cf. 103e) is only sufficient in fortunate cases like those of snow and cold, fire and heat, three and odd, where there is no doubt whether they can be called unqualifiedly (“always”) after their properties. In the case of the soul it would be a blatant *petitio* simply to assume that it can always be called alive. So for this special case, where there is neither physical nor conceptual necessity that the soul always should possess life, Plato introduced the extra criterion that it always imparts life to what it gets hold of. From the succession of the criteria, therefore, it seems clear, that only candidates which do satisfy the eponymy criterion would be subjected to the further test whether it “ἀεὶ ἐπιφέρει. . . .” For this is the way Plato proceeds; the question is nowhere whether the soul is alive because it brings life to the body, but only whether it is essentially

alive because it always does that (cf. 105d1; 3, 10).<sup>21</sup>

So it would seem that at least for our argument Plato can be acquitted of Hartman’s reproach that he confused “Pauline” with attribute-predication.<sup>22</sup> But the question remains whether Plato is really entitled to assume that the soul is an entity which can be called “alive” in a non-metaphorical sense. That he assumes this is also attested by a passage in *Republic* X, 608c ff. where Socrates offers another proof for the immortality of the soul and refers back to earlier arguments (611b-c) which can be only those of the *Phaedo*. In *Rep.* X Socrates argues that since everything can only be destroyed by its own congenital (“ἔμφυτον”) evil, but clearly human soul is not destroyed by its specific evil, vice, and the evils of the body cannot affect it, it cannot be destroyed at all. The weakness of this argument is apparent: there can be more than one specific evil (wood, e.g., can rot or be burnt, etc.) and things can also be destroyed in an “unnatural” way. Furthermore, vice may stand in relation to the soul just as ugliness stands to the body: it may disfigure but not destroy it. Yet, this does not exclude the possibility that nothing else could assail it. This passage confirms that Plato quite consciously conceived of the soul as of a substance with properties of its own, since he here compares it with other natural things such as wood, bronze, iron, grain, and the human body, which are all independent substances with properties and characteristic afflictions.

As to the exact nature of the soul we are left somehow in the dark by Plato in the *Phaedo* and also in *Republic* X.<sup>23</sup> The soul is sometimes treated as the principle of life (as in the Cyclical Argument), and the mind (in the Anamnesis Argument); in the Affinity Argument it fulfills both functions. In the argument in *Rep.* X the question is again left open. And this is the point where I want to venture my own criticism of Plato’s proof. So far, I have only pointed out that the proof, though formally correct, does not compel us to accept the immortality of the soul because one does not have to accept Plato’s concept of the soul as a separable substance, though one may do so. That Plato leaves the nature of the soul undefined, however, is a reproach from which one cannot, in my opinion, so easily release him, for this violates a rule which Socrates himself in several Platonic dialogues imposes on himself and on his partners: not to try to argue that a certain thing possesses a quality as long as one has not grasped the nature of the thing itself (cf. *Meno* 100b on “virtue”; *Rep.* I, 354c-e on “justice”; implicitly the same criticism is made at the end of the *Euthyphro* 15d and the *Laches* 199c ff).

It seems that this “Socratic” rule has been violated by Plato both in

the *Phaedo* and in *Republic* X, even though in the *Phaedo* Plato at the beginning of the Affinity Argument seems to promise such a clarification (78 b), and in *Republic* X he even admits his uncertainty about the nature of the soul. For he explains there (611a ff.) that in our present life our soul resembles the sea-god Glaucus, who is so battered, mutilated, and overgrown by maritime life that one cannot easily guess his real nature; thus for our soul "if it were raised out of the depth of this sea in which it is now sunk, and were cleansed and scraped free of the rocks and barnacles which, because it now feasts on earth cling to it in wild profusion of earthy and stony accretions by reason of these feastings that are accounted happy. And then one might see whether in its real nature it is manifold or single in its simplicity or what is the truth about it and how" (Shorey's translation). Since the soul's commerce with the earthly elements are so much stressed in the *Phaedo* too (cf. especially the affinity-argument 81b ff.), one wonders how, according to Plato, we are to know whether it is manifold or simple, separable or inseparable, as long as we do not know precisely what the soul is.

This "Socratic" criticism can also be formulated in Kantian terms, for it is this very knowledge, which seems to transcend the realm of our experience, of which Plato tries to convince us in his proof: that there is something in us which is a unity in itself, an immaterial entity which can be separated from the body and survives this separation in integrity and preservation of its faculties, an immortal soul.

About this, I assume, one could only be reassured if the "πρώτος πλοῦς" were possible, i.e., if we had direct knowledge of the causes of our generation and destruction in the way desired by Socrates in his discussion of the notion of cause (95c-99d), namely if we possessed knowledge of the fate of our soul before and after death; but Plato himself was aware that of this he could not give us an exact account but only a mythical description, the "likely story" with which Socrates concludes his talk (cf. 108d).

## Bibliography

Bluck, R. S. *A Translation of Plato's Phaedo with Introduction, Notes, and Appendices*. London: Routledge & Kegan-Paul, 1955.  
 Burnet, J. *Plato's Phaedo*. Oxford: Oxford University Press, 1911.  
 Erbse, H. "Philologische Anmerkungen zu Platons *Phaidon* 102a-107a." *Phronesis* (1969): 97-106.  
 Gallop, D. *Plato's Phaedo*. Translated with notes. Oxford: Clarendon, 1975.

Hackforth, R. *Plato's Phaedo*. Cambridge: Cambridge University Press, 1955.  
 Hartman, E. "Predication and Immortality in Plato's *Phaedo*." *Archiv für Geschichte der Philosophie* 54 (1972): 215-28.  
 Keyt, D. "The Fallacies in *Phaedo* 102a-107b." *Phronesis* 8 (1963): 167-72.  
 Loriaux, R. "Phédon 101d-e." *Zetesis Album Amicorum*. Antwerpen-Utrecht: E. de Strycker, 1973, 98-107.  
 O'Brien, D. The Last Argument of Plato's *Phaedo*." *Classical Quarterly* 17 (1967): 189-231 (I) and 18 (1968): 95-106 (II).  
 Scarrow, D. S. "Phaedo 106a-106e." *Philosophical Review* 70 (1961): 245-52.  
 Schiller, J. "Phaedo 104-105: Is the Soul a Form?" *Phronesis* 12 (1967): 50-58.  
 Vlastos, G. "Reasons and Causes in the *Phaedo*." *Philosophical Review* 78 (1969): 291-325.  
 —. "An Ambiguity in the *Sophist*." In *Platonic Studies*. Princeton: Princeton University Press, 1973. 270-322.  
 —. "The Unity of the Virtues in the *Protagoras*." In *Platonic Studies*. Princeton, N.J.: Princeton University Press 1973. 221-69.

## Notes

1. For a critical discussion of these arguments see Bluck 18ff. Bibliographical references are given at the end of this article.
2. If Plato seems to indicate some dissatisfaction here then it is not with the argument itself but with the fact that they had not pursued the "πρώτας ὑπόθεσις" as far up as possible. For the meaning of "ὑπόθεσις" here and in 101d3, see the article by R. Loriaux.
3. For a detailed elucidation of these difficulties, see the article by G. Vlastos "Reason and Cause in the *Phaedo*," esp. 308.
4. The vague meaning of "ἐναντίον" allows for quite different pairs of somehow opposed qualities. For the difficulties with relative opposites and their forms cf. Vlastos 315 n. 64.
5. As will be pointed out later, Plato may have had more important reasons for reminding the reader of the "law of opposites" at the beginning of the last argument.
6. That this distinction is new here seems to be indicated by the vocabulary—"τὸ ἔχον" and "τὸ ἔνον" (103 b) for the possessor and its quality; the term "quality" is only later introduced in *Theaetetus* (182 a). As we shall see, however, not just the novelty may account for the loose terminology; Plato may have wanted to be able to apply the distinction to cases as different as fire-heat, three-odd, etc.
7. For a discussion of the problem presented by the disparity of the examples cf. Gallop 199 ff.

8. In fact, it is in 106b3-4 presented as the *outcome* of the whole argument, not as a presupposition, that the soul will not be “τεθνηκυῖα.”

9. The condition “ὅταν περ ἦ” is then not a meaningless condition in the case of the soul, as O'Brien claims (I, 231) but necessary as long as step III is not established.

10. That Plato accepted self-predication for the forms in the *Phaedo* is suggested by 100c4-5 and 102d6 ff.

11. This would provide a more compelling reason for Socrates' recalling and clarifying the law, especially since Socrates' main partner, Cebes, declares that he had not needed it (102c).

12. Bluck often seems to put his finger on this very point (cf. 25; 119; 191 ff.); his explanation, however, that Plato tries to make the distinction between contrary and contradictory opposition passes over the most important point. The same assumption seems to underlie Scarrow's interpretation (cf. 24).

13. Cf. Keyt 169; O'Brien I, 216 ff., although O'Brien is partly critical of Hackforth's overall interpretation (cf. 217).

14. It seems significant that in *Parm.* (130c) fire, like water and man, is mentioned as one of the entities of which Socrates expresses doubt whether they *had* forms, not whether they *are* forms.

15. It would also seem that the “ἀδικού” and the “ἀμουσού” in 105d-e are not the immanent forms but the *things* that cannot accept the respective properties.

16. For a more thorough discussion of the meaning of the metaphor cf. Gallop 195 ff.

17. To the question whether the numbers, and in consequence the soul, are to be considered as forms, see the article by J. Schiller, esp. 57-58.

18. I leave aside the difficult question whether there also *is* a form of soul. Since Plato does mention the form of life itself (106d) it may well be that he also accepted a form of soul in which our individual souls participate.

19. For further elucidation of this notion see Vlastos, “An Ambiguity in the *Sophist*,” *Platonic Studies*, 270 ff; cf. also “The Unity of the Virtues in the *Protagoras*,” *Platonic Studies*, 221 ff, 152 ff.

20. Exceptions are, of course, the genuine cases of Pauline predication. Whether St. Paul was aware of the metaphorical character of his speech I do not know. For Plato it is only relevant that he did not infer “brings along → possesses.”

21. This point seems to have completely escaped the notice of D. Keyt, who accuses Plato of the fallacy of equivocation, that “athanatos” is at first used simply in the sense of “not dead” but later in that of “immortal.”

22. It should be stressed that this defense of Plato's argument does not affect the problem of self-predication of the forms.

23. Plato himself may already have seen problems which he attacked again in the *Phaedrus* and the *Timaeus*. For a more profound discussion of the nature of the soul in the *Phaedo* cf. Gallop, 88 ff.

## 13

### A Defense of Plato's Argument for the Immortality of the Soul at *Republic X* 608c-611a

Eric A. Brown

The immortal soul is one of the central elements of Platonism, inextricably linked to the account of learning as recollection (*anamnēsis*), the eternal Forms, and divine retribution for rights and wrongs. For this reason it is perhaps surprising how little attention gets paid to the immortality of the soul in the first nine and a half books of the *Republic*. But surprise typically gives way to sharp disappointment when we arrive at Plato's argument for immortality at 608c-611a. This argument has been almost unanimously scorned by commentators. Julia Annas proclaims that “this is one of the few really embarrassingly bad arguments in Plato,” and she refers to it as “a ridiculous little argument.” In his magisterial two-volume commentary on the *Republic*, the cautious James Adam identifies “difficulties” which are “serious, and possibly fatal.” And more recently, in his commentary on *Republic X*, Stephen Halliwell asserts that “it is clear, I think, that the beliefs or convictions which concern Plato at this point outrun the scope of cogently rational dialectic.” The message of gloom goes on and on. In despair, a charitable interpreter might decide to follow J. A. Stewart: “The argument for the immortality of the soul in *Republic* 608c-612a is formally so inconclusive that it is impossible to suppose Plato to be serious with it.”<sup>1</sup>

Don't believe the hype. Plato's argument for the immortality of the soul in *Republic X* is a fine argument. I do not mean that Plato manages to argue for his conclusion by a valid deduction from completely uncontroversial premises. (But then, when the conclusion is that the soul is

immortal, do we really expect that sort of argument?) Rather, he pulls together some interesting, intelligible, non-question-begging, and typically Platonic assumptions, and he constructs a valid argument from these. As I shall argue by a detailed analysis of 608c-611a, this argument, far from being embarrassing, builds in interesting and fruitful ways on the rest of the *Republic*. I proceed first by working through the passage and laying the argument bare, and then by considering the most pressing criticisms it faces. As an act of *mimesis*, I also offer in an appendix an analysis of the argument for reincarnation which Plato appends at 611a4-9.

## I

After defending the banishment of the poets in the first half of Book Ten, Plato turns to the rewards of virtue. He refers to the rewards discussed at the close of Book Nine (*eirēmenōn*, 608c3), and he has Socrates suggest that the conversation should move on to the “greatest” wages and prizes of virtue, those bestowed by gods and human beings in this life and the next (612a8 ad fin.).<sup>2</sup> But before we hear about these rewards, we must be sure that there is a next life: we must understand that the soul is immortal.

The immortality of the soul is introduced quite playfully, when Socrates apparently stuns Glaucon (608d3-12):

“Haven’t you perceived [*ēisthēsai*],” I said, “that our soul is immortal and never perishes?”

And he fixed his eyes on me with amazement and said, “By God, I have not—but *you* are able to say this?”

“If I don’t do an injustice [*adikō*], at least,” I said. “And I think you can, too, for it isn’t so tough.”

“For me it is,” he said. “But I would enjoy hearing from you this ‘not-so-tough’ thing.”

“You just might,” I said.

“Get on with it,” he said.<sup>3</sup>

Here we have Plato’s lighter side, and perhaps part of his portrayal of the historical Socrates, as well. There is certainly something a bit funny, although not hilarious, about the question “Haven’t you perceived that our soul is immortal?”, because the immortal soul is not something *perceived*. And there seems to be another inside joke lurking in Socrates’ odd way of allowing for error (“If I don’t do an injustice”).<sup>4</sup> Socrates’ odd locution ambiguously suggests that he and Glaucon can easily prove

that their souls do not perish only if they do not do an injustice, and this suggestion seems ironic once we realize that, according to the ensuing “not-so-tough” proof, their souls do not perish *even if* they do an injustice. Plato, it seems, is having fun.

But this fun has caused some commentators consternation. What is the point of jokes here?<sup>5</sup> The point is, I suggest, that it is difficult to argue for the immortality of the soul. Precisely this difficulty makes it funny that Socrates should be so casual and explains Glaucon’s surprise. Glaucon cannot be surprised by the mere *thesis* that the soul is immortal. After all, he is supposed to be well familiar with the Forms (475e6 ff., 596a5 ff.) and with Homer (Book Ten, *passim*), and it is extremely doubtful that a student of the Forms or of Homer could be utterly unfamiliar with the possibility of an immortal soul.<sup>6</sup> So, Glaucon’s surprise almost certainly highlights the difficulty of *proving* this thesis: Glaucon expects Socrates to have arguments for his positions and is stunned by Socrates’ blithe introduction of a tenet which seems so difficult to back up by argument. Like Simmias and Cebes in the *Phaedo*, he is skeptical of arguments on behalf of the immortality of the soul.

Having got our attention and signaled the difficulty of the argument, Plato then lets Socrates get it going (608d13-e5):

“Do you call something good,” I said, “and something evil?”—“Yes.”

“Do you understand these things the way I do?”—“How’s that?”

“The evil is entirely what destroys and corrupts, and the good is what preserves and benefits.”—“Yes,” he said.

Socrates starts out slowly and fixes the general direction of his argument. Although he is arguing for the immortality of the soul, his argument builds not on metaphysical considerations but on ethics. The *Republic*’s ethical theory will prove to be the foundation of the argument, and this slow start from broad ethical principles makes that clear right off. Socrates is nudging Glaucon, to keep him in mind of what has preceded in the *Republic*. We all, then, need to be in mind of how the *Republic* presents evils as corrupters and destroyers, and goods as preservers and benefitors.<sup>7</sup>

With Glaucon in the proper frame of mind, Socrates is ready to get to serious work (608e6-609a5):

“Well then, do you say that there is something evil and something good for each thing? Like eye disease for eyes, disease for the whole body, blight for corn, rot for wood, rust for bronze and iron, and, as I say, a natural evil and affliction [*nosēma*] for just about every single thing?”—“Yes,” he said.

Socrates is introducing the central concept for the argument, the natural evil (*sumphuton kakon*), and he gives us three points to consider.<sup>8</sup> First, he says that *almost* everything has a natural evil. Plato never says what things do not have a natural evil, but presumably, since natural evils will be defined as things which necessarily make their possessors bad and tend to destroy them, gods and Forms do not have natural evils.<sup>9</sup> Second, Socrates also suggests that nothing has more than one natural evil. Although this point is not offered with any great explicitness, we need to take it seriously, as it will turn out to be a crucial premise in the argument. Finally, he gives us some examples of natural evils. He has not told us exactly what a natural evil is, but these examples join the background concerns from the *Republic*'s ethical theory as important constraints on our understanding of that concept.

With these constraints in mind, let us go back to the text, for it adds more information about natural evils (609a6-8):

"Then whenever one of these is present in something, it makes the thing to which it is present bad [*ponérōn*], and in the end [*teleutōn*], it dissolves it entirely and destroys it?"—"Of course."

Natural evils possess two explicit *propria*: first, the natural evil of *y* makes *y* bad, and second, it dissolves and destroys *y*. We might take these *propria* to be claims about causal sufficiency: a natural evil is by definition (i.e., necessarily) causally sufficient to make its possessor bad and then to destroy it. But in advance of further evidence, we should not insist on this reading. Plato is unfolding his concept gradually, and while the rest of the argument will assume that all natural evils are causally sufficient to make their possessors worse, it will back away from the suggestion that all natural evils are causally sufficient to destroy their possessors.<sup>10</sup>

His characterization—making bad and dissolving—also suggests that natural evils work *on the inside*, causing decay and weakening a thing's constitution. In fact, throughout the argument, when Plato talks of destruction by evils, he uses words which suggest, although do not require, a model of internal decay. Just to look at the main stretch of the argument, we find frequent mentions of "corruption" (*diaphtheirein*) at 608e3, 609b1, d2, and d6, and "dissolution" (verbs in the family of *luein*) at 609a7, a10, b5, c2, and c6. (We might also compare "melt" [*tekein*] at c6 and "wither" [*marainein*] at d6, or notice that Plato refers to a "natural evil and affliction or disease" [*sumphuton kakon te kai nosēma*] at a3-4.) *Diaphtheirein* need not refer to destruction by means of internal decay,

but it does call to mind the internal moral corruption and decline portrayed in Books Eight and Nine, and all of these word choices together constitute a pattern suggesting that natural evils work internally to worsen and eventually to destroy.

Moreover, this suggestion is fully endorsed by our two constraints on interpretation. Socrates' examples—disease, rot, rust—represent *internal* corruption, decay, dissolution. The examples suggest that while evils in general may work in many ways, natural evils work from within. Of course, no reader of the *Republic* should be surprised by this. Throughout the work, we have been concerned with one particular evil, injustice, as a case of internal corruption and dissolution: Plato notoriously makes one's injustice a bad condition of one's own soul, and not one's bad treatment or poor consideration of others. So Glaucon, put in the mind of the *Republic*'s ethics, readily assents to this internal picture of natural evils: the natural evil of *y* makes *y* bad and destroys *y* by dissolving it.

It is important that we and Glaucon have this full background in mind, because Socrates is about to make a small inference from it (609a9-b3):

"so [*ara*] the natural evil and badness [*ponérēia*]<sup>11</sup> of each thing destroys it, or if this will not destroy it, then no further thing would corrupt-destroy [*diaphtheireien*] it. For [*gar*]<sup>12</sup> nothing good will ever destroy anything, nor again will something which is neither evil nor good."—"How could it?" he said.

Things which destroy are evil, and so, obviously enough, good things and neutral things do not destroy. But, Socrates is adding, not even evils can destroy a thing which has no natural evil. This is the central premise of the argument, what I shall call the *essential destructibility claim*. There are three problems in understanding this claim. First, it is apparently in tension with the characterization of natural evils which we just encountered. That characterization seemed to say that the natural evil of *y* is necessarily causally sufficient to destroy *y*, but the essential destructibility claim suggests that there might be natural evils which do not destroy. I suggested earlier that we should allow Plato time to unfold his concept, and now we see why. The central premise of the argument makes it clear that Plato is not committed to the claim that natural evils are necessarily causally sufficient to destroy their possessors.

Second, the essential destructibility claim is susceptible of many interpretations, and for this reason it has been the focus of much confusion about Plato's argument. He is clearly committed to the claim that if the natural evil of *y* does not destroy *y*, then *y* is indestructible. But there are

two general ways of understanding this claim. According to some, Plato holds that the natural evil of  $y$  is causally necessary for the destruction of  $y$ , whether as the sole cause, the primary cause, or a partial cause.<sup>13</sup> In other words, Plato is committed to something like one of these:

- A. Except for the natural evil of  $y$ , nothing can contribute at all to the destruction of  $y$ .
- B. Except for the natural evil of  $y$ , nothing can contribute to the destruction of  $y$  unless it brings about the natural evil of  $y$ .
- C. Except for the natural evil of  $y$ , nothing can contribute to the destruction of  $y$  unless it somehow works along with the natural evil of  $y$ .

But all of these interpretations leave Plato vulnerable to some easy counterexamples. Wood, for example, is easily destroyed without any causal role being played by its natural evil, rot, for perfectly healthy wood can be burned.<sup>14</sup> For this reason, many interpreters have thought Plato's argument to be a nonstarter, resting on a manifestly false, indeed silly, premise. But we need not give up so quickly.<sup>15</sup> For Plato's essential destructibility claim could be about logical, not causal, necessity. He might be making the logical point that any destructible thing must be capable of being destroyed by its natural evil:

- D. If there were no natural evil of  $y$  capable of destroying  $y$ , then nothing could destroy  $y$  (but if there is such a natural evil, then  $y$  can be destroyed by something other than the natural evil of  $y$ ).

Here, the claim is not that the natural evil of  $y$  must play some causal role in the destruction of  $y$ , but that it must be *able* to destroy  $y$ .<sup>16</sup> This interpretation is not subject to any stunning counterexamples, and I believe that it is exactly what Plato has in mind by his words at 609a10-b1.<sup>17</sup>

So the main difficulty would then seem to be a third one, namely, explaining why Plato should think that this central premise is true and why he should present it as an inference (*ara*, “so”—609a9). But even this difficulty melts away if we reflect on the background that Socrates has nudged into Glaucon's mind. The main argument of the *Republic* seeks to establish that internal, corrupting evils are far the most destructive dangers, so powerful that no amount of external rewards can make up for having an unjust soul. In fact, the *Republic* is consistently, emphatically concerned about internal health and decay, in contrast to external threats. Thus, the *Republic* does not say that a city is invulnerable to external attacks, but it does say that if that city is just and healthy, it has nothing to fear from those attacks. That is why when Kallipolis falls,

it will fall from within (546a1 ff.). On this background, it is natural to think that anything invulnerable to internal decay would be indestructible. Since these thoughts lie behind the articulation of natural evils, the essential destructibility claim is a plausible inference.

Furthermore, the grounds for the inference are entirely intelligible, and not unreasonable. According to the essential destructibility claim, if things are destructible, they must be destructible by some natural, internal process of decay and dissolution, and thus if something is not destructible by any natural, internal process of decay and dissolution, then it will not be destructible by anything external. Plato's point is that mortal things are mortal because of some feature of their own, internal constitutions, and if a thing should not have a mortal constitution, an essence subject to decay, then that entity has the essence of an immortal thing. In other words, mortality is a property of a thing's essence, and not a fact about how that thing happens to be situated relative to external features of the world. Against this, we might want to say that some things are mortal not on account of their own weakness, but at least in part on account of the surpassing strength of external threats. That is, we might insist, like Spinoza, that some things pass away even though their destruction is not part of their essence, or we might insist that mortality is a relational property. Such positions would provide a serious objection to Plato's argument, but it is far from obvious that Plato would have to bow to their force. It is in principle debatable whether mortality must rest in essential destructibility, as Plato holds, or in merely being weaker than externals, as opponents might hold. It is likely to be a messy debate, as the line between internal and external is not so easily kept bright and clear, but it is important to recognize that Plato is not done in from the start. His essential destructibility claim is intelligible, defensible, and most importantly for our interpretation, tightly bound to the ethical-cum-political theory of the *Republic*.

Plato now does what he often does in presenting his arguments: he takes stock and puts the main point(s) again (609b4-7):

“So if we should find some existing thing for which there is an evil which makes it bad but is unable to dissolve and destroy it, then won't we know then that there would be no destruction for something of such a nature?”—“This seems reasonable,” he said.

Socrates and Glaucon agree to a sort of decision procedure based on the essential destructibility claim. For any  $y$ , we look for the natural evil which makes it bad, and examine whether that special evil can destroy  $y$ . If it cannot, then, according to the essential destructibility claim,  $y$  is indestructible. The effectiveness of this decision procedure depends on the

assumption that each thing has at most one natural evil, but Plato is by no means sneaking this assumption into the argument. It has been present throughout and especially at the outset (608e6-609a4). The decision procedure introduces nothing new; it simply recasts the argument's main premises to frame what it is to come.

With the decision procedure in place, Socrates and Glaucon can consider the case where the human soul is *y*:

"Well then," I said, "doesn't the soul have something which makes it evil?"

"Indeed," he said, "all the things we were talking about just now, injustice, intemperance, courage, and ignorance."

Glaucon, as we have already noticed, is very much in mind of the *Republic*'s ethical theory, for he quickly produces the four cardinal vices as that which necessarily makes any soul to which it is present bad.

Since vice has been identified as the natural evil of the soul, we need only consider now whether or not vice can destroy the soul in order to know whether the soul is indestructible. Naturally, Plato has his interlocutors turn to this question next (609c2-d3):

"Then can any of these things dissolve and destroy it? And keep your head lest we be deceived by thinking that an unjust and foolish person, whenever he is caught doing an injustice, is then destroyed by his injustice, which is a badness of the soul. Rather, think of it this way: just as the badness of body, which is disease, melts the body and destroys it and leads it to being a not-body, so too all the things we mentioned just now are brought to not-being by their own corrupting vice, by its attachment and presence in them—isn't this so?"—"Certainly."

These words carefully demarcate the climactic question. Plato is careful to distinguish how injustice might destroy a person who is caught and killed for his injustice from how injustice would destroy a person *as a natural evil*. Based on our understanding of natural evils, this distinction is immediately intelligible: the first sort of destruction is from without while natural evils destroy from within.

Having issued his clarificatory *prolegomena*, Socrates can finally put the important question (609d4-8):

"Come then, and look at the soul in the same way. Does injustice, by being in it, and does the rest of vice, by being within and attached, destroy and wither it, until, having led it to death, it separates it from the body?"—"Never," he said, "does *this* happen."

Critics have been mighty unhappy about this exchange, and we shall come to consider their complaints in due course. For now, for purposes of exposition and interpretation, we need to see this exchange as Plato does, as the triumphant answer he needs. Glaucon answers that the soul's natural evil does not destroy the soul, and this answer, taken together with the essential destructibility claim, gets Plato the conclusion that the soul is indestructible. Plato does not make this conclusion explicit until 610e5, when he moves to cap the argument formally, but he nonetheless proceeds with full recognition of having essentially reached his conclusion. He immediately has Socrates shift the focus of the conversation slightly: Socrates turns from laying out the argument to considering and challenging the opposition. Throughout this challenge to the opposition, it is clear that Socrates knows his argument is virtually done (see especially 610a8-9), and when he comes to giving the argument its final conclusion, Socrates proceeds quickly and formally, without any substantively new premises (610e10-611a2).

Thus, the real construction of the argument is complete, but before Plato caps it off, he wants to explain and strengthen his position by considering the opposition. This maneuver is familiar; it is just as we today expect a decent philosophical argument to consider the opposition and its possible objections. Furthermore, just as we frequently expect, Plato takes the opportunity of addressing the opposition to issue a challenge and to shift the burden of proof.

To make sense of this stretch of the argument (609d9-610e9), we need to recognize who Plato's opposition is. The opposition obviously must hold that there is no immortal soul, but this position could be defended in many ways. One might, for example, disagree with the soul-body dualism which is required by the immortality of the soul, or one might accept the dualism and yet suppose that the soul perishes at some point. Plato targets those who take the soul to be destroyed when it separates from the body. He does not concern himself with any opponents who deny a soul-body dualism, nor does he consider opponents who imagine the destruction of the soul at some time other than its separation from the body. But Plato does not choose his target arbitrarily. In the *Phaedo*, it is repeatedly said that the soul's destruction at separation is a view or fear of the many (69e5-70a7, 77b3-6, 80d8-e1). Thus, in the *Republic* Plato is contrasting his own argument for the immortality of the soul with the popular view that the soul might perish when a person dies.

Plato's strategy for dispensing with this opposition is quite general. Since he takes it as established that if anything at all can destroy the soul, its own natural evil must be able to, he goes on to challenge anyone proposing an agent of the soul's destruction to show that that agent brings

about vice. For if that agent does not bring about vice, then vice cannot be said to destroy the soul, and if vice cannot be said to destroy the soul, then by the essential destructibility claim, the soul is indestructible. Until it can be shown that vice destroys the soul, the opposition's candidates for destructive agents are rejected by the terms of the argument.

This general strategy assumes that we stick closely to our understanding of how a natural evil corrupts, dissolves, and destroys. We must not say that vice destroys when we are caught doing vicious things and put to death, for that is not a destruction brought about by vice in its corrupting, dissolving way. Plato made this point earlier (609c2-5), and now he refers back to it (610d3-4). Plato's Socrates keeps us to the terms of the argument and the ethical theory on which it depends in order to press Glaucon away from the popular fear and toward the logical implication of his premises.

The challenge issued by this general strategy proves to be quite successful. Because the destruction of the soul is popularly linked to its separation from the body, the agents which are commonly thought to destroy the soul are those things which kill people, that is, cause the separation of soul from body. But these agents, such as disease, can hardly be said to cause a person to become unjust or generally vicious, and they kill virtuous people just as readily as vicious ones. These agents thus cannot be said to be the soul's natural evil nor can they be said to bring about the soul's natural evil. So these agents do not explain how the soul has a natural evil which destroys it, and this explanation is what must be provided by anyone seeking to show that the soul is not indestructible. The premises of Plato's argument put an enormous burden on the popular view and thereby lend support to his own position.

Having wielded his argument against the popular view, Plato allows Socrates to provide the argument with its formal conclusion (610e10-611a3). He writes,

"Thus when something is not destroyed by any evil, whether its own or another's, it clearly must always exist, and if it always exists, then it is immortal."

"Necessarily," he said.

This is quick and simple, and we can readily fill in its meaning for the argument at hand. Since the soul is indestructible, it always exists, and since it always exists, it is immortal. *Quod erat demonstrandum.*

## II

Criticism has been launched at three primary targets in this argument. Commentators have quarreled with the essential destructibility claim, with Glaucon's answer that vice does not ever destroy the soul, and with the implicit metaphysical assumptions built into the argument. At each of these points, some readers have thought Plato to be guilty of begging the question, and most readers have thought Plato to be moving far too fast. I happily concede that the essential destructibility claim, Glaucon's answer, and Plato's metaphysics are all very arguable and that the argument as a whole is therefore unlikely to be persuasive. But I take it that these concessions do not impugn Plato's argument; we should not expect an argument for the immortality of the soul to be unarguable or absolutely compelling. Rather, to impugn Plato's argument, we need to show that he has begged the question, or has used some easily falsifiable premise, or has unreasonably provided no account for another. I argue that he is innocent of these charges. At each of the most arguable junctures of his argument, Plato proceeds intelligibly and reasonably, making ingenious use of assumptions which are pervasive in the *Republic*. I argue for this conclusion by considering six criticisms, two for each of the three most debatable features of Plato's argument.

Julia Annas crisply states the first criticism: "the argument only works if each thing has *one* peculiar evil; otherwise we could not argue that if justice [sic] does not destroy the soul, nothing else will. Could not the soul have another natural peculiar evil besides injustice?"<sup>18</sup> This is perfectly correct. Plato does indeed commit himself to there being at most one natural evil for each thing. In other words, there is at most one evil for *y* which is necessarily causally sufficient to make *y* worse. So Annas is right to wonder whether the soul could not have additional potential natural evils. We could argue with Plato about this question, but we should notice that our argument will be with the full ethical theory of the *Republic*, and not just with a separable argument about the soul in Book Ten. For it is a pervasive assumption in the *Republic* that each thing has one function, one excellence which consists in performing the function well, and one vice which consists in performing that function poorly, and it is a pervasive assumption that virtue is the particular excellence of the soul, while vice is the particular evil. Plato cannot just give up the idea that each thing has at most one potential natural evil.<sup>19</sup>

This first criticism of Plato's essential destructibility claim should not be confused with a second one, which takes the claim to be that any given thing can only be destroyed by one agent. If Plato were saying that, he would be saying something unarguably false, and thus he would in-

deed be deserving of censure. But we have already encountered this criticism, and we have already seen that it rests on an unnecessarily uncharitable reading. On any reasonable interpretation of the essential destructibility claim, this second criticism just misses its target.

We encounter more difficult problems when we consider Glaucon's answer that vice does not destroy the soul. First, we might worry that Glaucon is confused by Socrates' question, and that his answer depends on his confusion. Recall that Socrates asks (609d6-7), "Does injustice, by being in the soul, and does the rest of vice, by being within and attached, destroy and wither it, until, having led it to death, it separates it from the body?" J. A. Stewart would find the roots of an equivocation here: "The argument is, that, as Injustice, the proper vice of the Soul, does not cause "Death," in the sense of the separation of Soul from body, nothing else can ever cause "Death," now, however, to be understood in the sense of the annihilation of the disembodied Soul itself."<sup>20</sup> By Stewart's lights, Plato just conflates the soul's destruction (the death of the soul) with its separation from the body (the death of the embodied person). If this is right, then Glaucon's answer might depend on this unfortunate conflation. Yet we need not read the exchange in this way, and while I believe that Stewart is on to something, I do not think that he has found an equivocation. Rather, we can and should understand Plato to associate the soul's destruction with its separation from the body without identifying the soul's destruction and its separation. That is, Plato assumes that if the soul were to be destroyed, it would have to be destroyed at the time when it separates from the body, but he never simply conflates the death of the soul with the death of the embodied person.

Perhaps someone will now tag in for Stewart and object that Plato is assuming too much. Why should the soul have to be destroyed only at the time when it separates from the body? Plato has two good reasons for making this assumption. The first applies the principle of sufficient reason. If the soul is to be destroyed, we need to ask why it should be destroyed at this time rather than that. We might well think that the only reasonable candidate time is at the separation of the soul from the body. On this reasoning, if the soul is destructible, it must be destroyed when it is separated from the body. Second, even if we blanch at this reasoning, we need to recall that Plato's principal opposition accepted it. Plato is defending the immortality of the soul against the view that the soul is destroyed when it separates from the body. So Plato has good enough reason to suppose that if the soul is destroyed, it is destroyed when it is separated from the body, and he need not be taken to be supposing that the destruction of the soul is simply identical to its separation from the body. Stewart's criticism is disarmed.

Others who are troubled by this same crucial step in the argument have worried about Glaucon's grounds for thinking that vice does not destroy the soul. Adam puts this fourth worry especially well: "Is [this] intended as an appeal to experience? Even if we allowed that experience is the proper tribunal, our experience of the effect of injustice on a human soul is limited to a single life; and why should not one soul wear out many bodies and perish at last through its own vice?"<sup>21</sup> Adam rightly realizes that Glaucon's answer rests on some empirical evidence. Indeed, if the answer were given on *a priori* grounds, it would be difficult to acquit Plato of begging the question.<sup>22</sup> But Adam is skeptical about Glaucon's evidence, for two separable reasons. First, Glaucon is making a claim about the soul's destructibility which cannot rest on direct experience, for we never directly perceive the soul and its fate after separation from the body. (We are reminded of the humorous exchange which precedes the argument: Glaucon does not *perceive* that the soul is immortal.) Glaucon's answer must rest entirely on indirect evidence. Furthermore, Glaucon's evidence might be, for all we know, entirely too indirect. The soul might perish not directly upon separation from the body, but only after several lifetimes, Adam suggests, and Glaucon's experiences in this single life are entirely inadequate to deny this possibility. So, Adam suggests, Glaucon's crucial answer rests on a paucity of evidence.

Glaucon's response, however, is far from utterly groundless. Plato has reasons to think that Glaucon's indirect evidence is adequately relevant to deciding whether vice can destroy the soul. First, we can address the second half of Adam's critique. We should notice that Glaucon answers with a negative: the soul is never seen to perish on account of vice. Establishing a negative through any appeal to indirect evidence is more perilous than establishing an affirmative; to be thorough, Glaucon would have to consider every possible scenario of the soul's destruction on account of vice and consult any possible evidence for that scenario. In the face of such a far-reaching demand, Plato's argument adopts a reasonable, heuristic strategy: it considers the most plausible, most widely accepted scenario. If that scenario for the destruction of the soul fails to harmonize with the available evidence, then we have some reason for thinking that no such account of the destruction of the soul should be accepted. When we follow such a procedure, our conclusion is not iron-clad, but we have not begged the question or proceeded without reason.

Plato thus assumes for the sake of argument that the soul perishes as soon as it separates from the body, and this assumption reasonably represents a more plausible rival than Adam's alternative that the soul perishes only after several lifetimes. We must resist the temptation to entertain Adam's alternative on the grounds that the soul is destructible only

after it is itself worn out by many lifetimes spent embodied. On this line, the soul's natural evil cannot destroy the soul: only a combination of vice and weariness can destroy the soul. This violates the essential destructibility claim's insistence that a destructible thing's natural evil must be able to destroy it. As we have seen, Plato holds that mortality is a property of a thing's essence, and not a relational property. Thus, if the soul is destructible, it must be destructible at least primarily from within. If the soul is destructible from within, a given soul might in fact be destroyed from without or by a combination of factors from within and without. At least some souls, however, must clearly be destroyed at least primarily from within, and this possibility is not explained by the scenario in which only *wearied* souls perish from their vice.

Without this line, however, it is harder to see what plausibility recommends Adam's alternative. And the alternative has its own problems: Why does vice worsen the soul immediately, but destroy it only over much time? Does the cumulative effect of vice require that a reincarnated soul begin its new life marred by the vice of its previous life? If not, how is the dissolution of the soul supposed to worsen in a secondary lifetime? For these reasons, Plato is right to stick with the more plausible, more widely accepted rival view of the soul's destruction and to ignore Adam's alternative. He is quite reasonably providing a focus for Glaucon's consideration where a full consideration of all the possibilities is not feasible. Plato's strategy leaves the door open for potential challenges and disputes, but it does not proceed unintelligibly or overhastily.

Having seen how and why Plato limits Glaucon's perspective to the assumption that the soul will perish at its separation from the body if it perishes at all, we are now in a position to evaluate Glaucon's appeal to indirect evidence for the soul's destructibility and address the first of Adam's two concerns. Because the most plausible account of the soul's destruction ties this event to the soul's separation from the body, our experience about what induces the soul's separation from the body is in fact relevant to our judgments about the soul's possible destruction. Glaucon would not be able to perceive vice destroying the soul directly, but he can perceive whether vice brings about the soul's separation from the body. And he notices what is obvious: vice does not seem to kill the vicious. Nor indeed does vice seem to bring the vicious closer to death, judging from the fact that many vicious people seem to be more lively than many virtuous people (610d7-e4). Vice seems entirely unconnected with the separation of the soul from the body. But if vice were to destroy the soul, we should expect to see vicious people brought closer to the destruction of their soul. We should expect, then, to see them approach their death, when the vicious soul separates from the body and then per-

ishes. But we do not see this. In fact, our best evidence points in the other direction.

There is no doubt that Glaucon's answer must extrapolate from limited experience, and there is no doubt that his answer depends for its persuasiveness on assuming the association between the destruction of the soul and the soul's separation from the body.<sup>23</sup> Thus, Adam is right to find room here for disagreement. I do not hope to have shown that Glaucon's answer demands our support. Rather, I hope to have shown that it does not beg the question or arrive entirely without warrant. Plato has at least revealed that conventional explanations of the soul's mortality fit badly with the demands of the essential destructibility claim. Conventionally, the soul perishes *because* it is separated from the body, but this story fails to explain how the soul is essentially mortal because it leaves the soul's natural evil out of account. If we are committed to the essential destructibility claim, then we have to ask whether vice can destroy the soul. At this point, we can at least say that no evidence seems to suggest that vice can destroy the soul, and that some evidence actually tells against this suggestion. This may not persuade us, but we ought at least to recognize its cleverness.

The fifth criticism again charges Plato with a *petitio principii*, this time by way of an implicit metaphysical assumption. Annas states this objection most clearly and trenchantly:

What is to prevent the relation between body and soul being of this [physicalist] kind, destruction of the body carrying with it destruction of the soul? Plato insists vehemently that this is not so (610a-b), but only by assuming what is never argued for, that the soul is a different kind of thing from the body, and distinct from it in being unaffected by what happens to the body. . . . As in Book I, we find a snappy little argument for a controversial conclusion depending on a question-begging premise which is never argued for.<sup>24</sup>

Of course, Annas is right to notice that Plato assumes soul-body dualism in this argument. This is another point at which we might reasonably choose to disagree with Plato and reject his conclusion. Right as this is, it is wrong to think that this dualist assumption begs the question. One could well hold that soul and body are two different sorts of thing and still maintain that the soul is mortal. Indeed, that is exactly the view that Plato attributes to most people in the *Phaedo*, and exactly the view that he targets in this *Republic* argument. Plato's assumption of soul-body dualism is necessary for his argument, but it is hardly sufficient.

Furthermore, it is not clear whether we should reasonably expect Plato to defend his assumption of dualism in this context. No doubt, if he

were pressed on the matter, he should. Or if he were making a full metaphysical inquiry into the soul, he should.<sup>25</sup> But neither of these conditions obtains. Plato is making an ethical argument for the immortality of the soul—an argument which appeals mainly to ethical claims and an argument which is designed mainly for ethical purposes—and this explains the absence of justifications for dualism. R. L. Nettleship recognized this point almost a century ago:

Plato makes no attempt here to deal completely with the question of the immortality of the soul. In the *Phaedo* he treats the whole question, but here it is only its bearing on morality that concerns him, and the question is touched upon just far enough to give completion to his picture of the destiny of the soul on earth. The point of view from which he argues that the soul is immortal is one which is in keeping with the whole subject of the *Republic*.<sup>26</sup>

Dualism and the metaphysics of soul are just not on the table here. Even when Plato follows up his arguments about immortality with a discussion of the true nature of the soul, he moves quickly and metaphorically, and he explicitly leaves open the crucial question of whether the true soul is simple or tripartite (esp. 612a3-6). Book Ten in general and this argument in particular are not intended to answer all the metaphysical questions, and if we dismiss the argument for just that reason, then we are failing to evaluate Plato's argument on his terms.

Finally, some people have thought that Plato is sneaky about assuming that there is a plurality of souls. Again, Annas puts the point clearly: "Even if the argument worked, why does it show not only that *the soul* is immortal but that there is a plurality of immortal souls? . . . One might take the argument to show only the immortality of *soul* thought of as something present in different places but not divided into separate individual souls (as though "soul" were a mass term)."<sup>27</sup> I do not deny that one might take the argument that way, but I think it obvious that Plato does not take it that way, and I have no trouble granting to Plato the assumption that there are a plurality of individual souls. This assumption, like that of dualism, remains implicit and undefended in this context, but once again, Plato's purposes in the *Republic* justify this omission. For the entire *Republic* has assumed that there are a plurality of individual souls: the tyrant is unjust not because of the conflicts within his share of soul-stuff, but because *his own* soul is in a bad way, and the philosopher is just not because of the relationships between the parts of her share of soul-stuff (whatever that would look like), but because the parts of *her own* soul are in agreement. This pervasive assumption is part and parcel of an arguable metaphysics, but as we have seen, Plato is not concerned

to spell out and defend the metaphysics of the soul in the *Republic*. His discussion of the soul in the *Republic* has other aims.

These last two criticisms, concerning Plato's metaphysical assumptions, prompt some general reflections on interpreting the Book Ten argument for the immortality of the soul. In any argument that Plato or anyone else makes, certain assumptions must be made without argument, whether implicitly or explicitly. In any argument for a controversial conclusion, some of these assumptions are likely to be arguable. We might reasonably ask the author of such an argument to defend her most arguable assumptions. But in certain times and places, that author might reasonably beg off defending one or another of these. "Such a defense will take us too far from our topic here, and I will have to take it up elsewhere," she might say. Or even: "I am not here addressing those who disagree with that point, for this discussion is aimed at those who have followed me sympathetically up until now." With the *Republic*, we are faced with a mute text, words unable to respond directly to our questions.<sup>28</sup> But it is not hard to see that Plato could make either of the two responses just imagined. For on the one hand, the argument for immortality stays close to ethics, avoiding metaphysics, and on the other, it presupposes the acceptance of the main principles of the *Republic*. In both ways, this argument continues the pattern set in Book Ten by the defense of the banishment of poets: both passages pick up loose ends and tie them from within the framework of the rest of the work.<sup>29</sup> This should warn us against taking the argument for the immortality of the soul as a complete, freestanding attempt to demonstrate its thesis.

In part because the argument for the immortality of the soul is not a complete, freestanding offering, no skeptic is likely to find it persuasive. But if I am right, it is not intended to persuade a skeptic. It is aimed at those who share certain metaphysical assumptions about the soul which the *Republic* is not concerned to defend and who are deeply committed to the ethical theory which has unfolded over the course of the *Republic*. Hence, we might reject Plato's essential destructibility claim because it fails to treat mortality as a relational property, or we might balk at his insistence that each thing has at most one particular excellence and one particular fault. We might question the grounds for Glaucon's crucial answer, especially in light of our skepticism concerning the essential destructibility claim. Or we might simply bring entirely different metaphysical assumptions to the table. But we ought not let our philosophical differences with Plato cloud our interpretive judgment of his achievement. Plato is arguing for a difficult thesis, and his argument contains no fallacies, rests on no unarguably false premises, and ingeniously builds on familiar theses from the rest of the *Republic*. Understood from within

the *Republic*—and not as an independent, freestanding metaphysical argument—Plato's effort deserves credit.

### Appendix: The Argument at 611a4-9

Having shown that the soul is immortal, Plato appends a brief argument to show that there is a finite number of immortal souls in a continuous cycle of reincarnations. Immediately after Glaucon agrees that souls are necessarily immortal, Socrates extends the conclusion thus:

"Let this be so, then," I said, "and if it is so, then you understand that the same souls will always exist. For they could never become in any way fewer, since none is destroyed, nor could they ever become more. For if anything immortal should become more, you know that it would become more from the mortal and everything would, in the end, be immortal.—'You speak truly.'

In the quickest of moves, Plato is repeating an argument from the *Phaedo* (72a-b). He does not need this conclusion in order to restore the rewards of the afterlife to the defense of justice: the immortality of the soul alone provides for the possibility of an afterlife with rewards and punishments. This addendum merely seeks to provide some grounds for believing that the afterlife looks something like the way the Myth of Er describes it. Here he tries to argue that there is a finite number of souls in a continual cycle of incarnations and afterlives, and this argument, if it persuades us, will make us more likely to accept the story attributed to Er. But this argument is considerably weaker than its predecessor, and Glaucon gives Socrates more credit than he deserves.

I shall offer two possible reconstructions of this argument. They are not mutually exclusive, but they take the argument quite differently. When I originally presented the first reconstruction, I learned that it had been offered by Proclus, in his pseudocommentary on the *Republic*, and so I shall call it the Proclan version of the argument. The second reconstruction takes Plato to be making a point much like one which Melissus of Samos at one point makes, and so I shall call that the Melissan version of the argument.<sup>30</sup>

First, the Proclan version. To understand the argument, Proclus suggests, we need to grasp the fact that Plato is assuming that there will always be mortal living things, and thus, there will always be things with souls. (For Plato as for any Greek, this "thus" represents a trivial inference, for to be animate is to have a soul.) The infinite future, then, con-

tains an infinite number of living things requiring souls. To explain the existence of all these souls, we have two options: either each living thing has its own unique soul, or a finite number of souls is continually recycled. Plato is arguing for the second option by reducing the first option to absurdity. The key moves, then, go like this. If each of the infinite future souls were to have its own unique soul, then there must be an indefinite supply of new souls, and this supply must either come from nothing, from other immortal things, or from mortal things. Plato assumes that only the third possibility should be entertained: the first option is presumably impossible on the grounds that *nihil ex nihilo fit*, assumed for Parmenidean reasons, if nothing else; and the second option is also disagreeable, because immortal things just do not spawn new souls and because there is no indefinite supply of immortal things to become souls by transmutation. But the third possibility creates difficulties as well: if an indefinite supply of souls is to be created from a finite supply of mortal stuffs, eventually everything will be immortal, and this situation contradicts the original assumption that there will always be a world with mortal living things. Hence, the supposition that each living thing has its own unique soul is reduced to absurdity, and we are forced to recognize that there is a finite number of souls continually being recycled. In this way, Proclus' reconstruction offers Plato a reasonable argument for a difficult conclusion.

This argument is not exactly watertight, however. In particular, it is quite unclear why immortal things cannot spawn new souls. This assumption is central, but entirely ungrounded. Furthermore, it is difficult to forgive Plato for skipping the metaphysical niceties in this particular argument, because in stark contrast to the main argument, which builds on ethical premises from the *Republic*, this appended argument offers nothing but metaphysical assumptions. Nonetheless, one criticism which has been launched against Plato's appended argument fails to stick to the Proclan reconstruction. Annas has complained that our common sense is not enough to assume that it is *absurd* that all mortal things should end up immortal.<sup>31</sup> On the Proclan reconstruction, though, this proposition is straightforwardly absurd, for it contradicts an assumption made earlier in the argument. So Annas' worry need not afflict us. And while this reconstructed argument is more tenuous than its predecessor, it is hardly unintelligible, nor is it obviously question-begging. It is simply incomplete, leaving unstated and unexplained a central claim.

Yet Proclus' clever interpretation seems to distort Plato's actual words a bit. For Plato seems to assume *as a completely general principle* that if some set of immortals comes to be larger, then everything will end up being immortal. Proclus understands this not as a completely general

principle, but as a specific claim against the backdrop of specific assumptions of indefinitely increasing numbers of souls. To be truer to Plato's words, we should seek a way of making sense of this argument which turns on that completely general principle. The argument in this case would go as follows:

- (1) (from the previous argument) The soul is immortal.
- (2) (from 1) The number of souls can never be made fewer.
- (3) (on grounds similar to those assayed above) If the number of immortal things is increased, it must be from the mortal.
- (4) If immortal things are increased, then everything eventually will be immortal.
- (5) It is absurd that everything should be immortal.
- (6) (from 4 and 5) The number of souls does not increase.
- (7) (from 2 and 6) The number of souls remains the same.

On this reconstruction of the argument, the same problem we found above about spawning immortals remains a concern, buried beneath premise (3), but now we also have problems motivating premises (4) and (5).

For the fourth premise, David Sedley has suggested that Plato has a Melissian point in mind. Melissus characterizes the One:

In this way, then, it is eternal and infinite and one and all alike. And it neither is destroyed at all nor does it grow larger nor does it become rearranged nor does it become sick or grieve, for if it were to suffer any of these things, it would no longer be one. For if it alters, what is must be not alike; rather, what was earlier is destroyed and what is not comes to be. In fact, if it were to become different by a single hair in a thousand years, it will entirely be destroyed in the whole of time.<sup>32</sup>

Without attempting to make sense of Melissus' entire point here, let us focus on the last sentence. This might be taken to mean that if a principle of change is admitted at all for what exists, then eventually—because the whole of time is, after all, a very long time—what exists will be entirely other than what it is now. Such reasoning might well explain Plato's general prohibition on mortal things becoming immortal. If we admit the principle that the finite number of mortal things can become immortal, then eventually, all of the things which are now mortal will have become immortal. And of course, since immortal things cannot, by definition, become mortal—for in that case they would not have been immortal—then if we admit the principle that mortal things can become im-

mortal, then we commit ourselves to the eventuality that there will be no mortal things.

This reasoning is not unproblematic, as it amounts to a sort of snowball argument. That is, just as a snowball which is rolled down a snowy hill gathers mass and speed, so too one change now will lead irrevocably to more and more changes. Why should the changes not be limited? Why should *any* increase of immortal things necessarily become a sufficiently large enough increase to exhaust the supply of mortal things? Perhaps Plato would respond by saying that nothing prevents an increase from becoming an exhaustive increase, but this response is not quite enough. I, for one, would be happier if Plato were to rest on the claim that mortal things just cannot *become* immortal: either a thing has a mortal essence or an immortal one, and there is no convertibility. (This thesis could be close to his grounds for the essential destructibility claim.) But Plato explicitly does not invoke this thesis. Instead, he suggests that if there were convertibility, it would go too far. And while it is true that the whole of time takes a while, it is not clear why the conversion of mortal things to immortal things must continue until there is nothing mortal.

Furthermore, on the Melissian construal of Plato's argument for the cyclical incarnations of souls, Annas' criticism holds. Why, on this version of the argument, should there always be mortal things, so that premise (5) is true? (Note that we cannot say, "Because there must be mortal stuff for the souls to be reborn into." That answer begs the question.) It is possible that Plato has resources elsewhere to bolster this claim. But as things stand, this subsidiary argument concerning the immortal soul is far more problematic than the prior argument for the thesis that the soul is immortal.<sup>33</sup>

## Notes

1. Julia Annas, *Introduction to Plato's Republic* (Oxford: Oxford University Press, 1981), 344; James Adam, *The Republic of Plato*, edited, with critical notes, commentary, and appendices, 2 vols. (Cambridge: Cambridge University Press, 1902), 423; Stephen Halliwell, *Plato: Republic 10* (Warminster: Aris & Phillips, 1988), 17; J. A. Stewart, *The Myths of Plato* (New York: Macmillan, 1905), 73n. Both Stewart (73) and Halliwell (17) decide that the failure of the argument is intended to prompt and explain Plato's subsequent turn to myth.

2. There is irony and subtlety in Plato's choice of words in this transition. In particular, it is significant that Plato uses the words "wages" (*epicheira*) and "prizes" (*athla*) for the "greatest" rewards bestowed by the gods and fellow human beings (608c1-2). Wages and prizes are precisely the ends valued most of all by the lower two parts of the soul, the money-loving appetitive part and the

victory-loving spirited part (see esp. 580d ff.). By choosing these terms, Plato makes two interesting points. First, he asks us to question whether the rewards bestowed by gods and humans are in fact the greatest rewards. Books Eight and Nine have disparaged wages and prizes as ends, and we have just been reminded of this disparagement at 608b6. So, by calling the appetitive and spirited ends the “greatest” rewards, Plato is using irony to remind us that virtue—the characteristic end of the rational part of the soul—is its own greatest reward. His choice of words tells us that while the restoration of rewards bestowed by gods and humans might complete the argument of the *Republic*, it nonetheless leaves unchanged what really matters. Furthermore, the choice of these terms offers us a new insight into the challenge put to Socrates by Glaucon and Adeimantus in Book Two. They had asked for a defense of justice by itself, minus rewards, and Socrates now tells us that they were, in effect, asking for an assessment of the value of justice as the rational part of the soul sees it, minus the value of justice as the other two parts of the soul see it. They had suggested that most people try to appear just only to reap wages and prizes, but once this picture is recast, Glaucon and Adeimantus should see how badly this sort of pursuit misconstrues value. Plato underscores these subtle, important points by consistent and careful control of his vocabulary concerning the rewards bestowed by gods and humans: hence, we find “prizes” (*athla* or *nikêtéria*) at 612d6, 613b6, 613c6, and 614a1 (and cf. 613b10-c4), and “wages” (*epicheira* or *misthoi*) at 612b8 and 614a1.

3. I am following Burnet’s Oxford Classical Text, and my translations are often based on Halliwell’s. All parenthetical references are to the *Republic*, unless otherwise specified.

4. I owe this point to Myles Burnyeat.

5. See, e.g., Annas, *Introduction*, 344.

6. I do not mean to imply that Homer presents anything like Platonically immortal souls, but only that someone like Glaucon should have been led to reflect about the possibilities of immortality by Homer’s depictions of life after death.

7. We might take 608e3-5 as the first premise of the argument, but I do not think that we should. If we do, we will most likely read it as something like, “*x* is evil if and only if *x* destroys and corrupts.” (I take it that the “entirely” (*pan*) licenses at least a biconditional.) But this claim does not fit the rest of the argument well, first because it comes in for revision—according to the argument, not every evil destroys—and second because the argument nowhere uses this claim as grounds for an inference. It is better, then, to take this passage as background to get Glaucon going than as the first premise.

8. Plato uses many different terms for this concept in addition to *sumphuton kakon* (609a3, a9) and *emphuton kakon* (610a2). Sometimes he refers only to *kakon* where the context makes it clear that we are talking about this concept (e.g., 609b5); sometimes he uses *oikeion* or *allotriion kakon* (610a2, a7, b6, e6); and sometimes *ponéria* (609a9, c5, c6, d10, e2, e4, 610a6, a7), *mochteria* (609e5), and even *kakia* (609d1).

9. I leave aside puzzles about the self-predication of Forms of bad things: does the Form of a natural evil cause itself to become bad? I take it that however

we feel about these puzzles, we can recognize that Plato seems committed to the claim that Forms *qua* Forms are not the sort of things to have a natural evil.

10. Elizabeth Asmis has suggested to me that 609a6-8 might not even apparently commit Plato to the claim that the natural evil of *y* is (necessarily) causally sufficient for the destruction of *y*. She suggests that we read the word *teleutón* in 609a7 as a conditional participle, so that Plato is saying something like, “it makes the thing to which it is present bad, and if it accomplishes its end, it dissolves it entirely and destroys it.” I find this reading a bit forced (see LSJ, s.v. *teleutaō* II), and so I prefer to suggest that Plato is gradually refining his concept of the natural evil before he puts it to use at 609b9 ff.

11. Roger Shiner suggests that Plato’s word choice here might be meant to echo the claim of 609a6-8 that a natural evil makes its thing “bad [*ponérōn*],” in which case perhaps we should understand Plato to be saying, “the natural evil and the badness [the evil induces].” In fact, I think that Plato is just introducing a new term for his concept of the natural evil, as he will use *ponéria* many times in this way (609a9, c5, c6, d10, e2, e4, 610a6, a7). Shiner is right, however, to draw our attention to the connection between the two passages: Plato readily associates the natural evil with its product, badness. This explains not only the use of *ponéria*, but also the use of *mochthéria* (cf. 609e5 with 609b5). It also gives us reason to take seriously the claim that natural evils necessarily are causally sufficient to make their possessors bad (609a6-8).

12. For concerns about this “for [*gar*],” see n. 17.

13. Adam, *ad loc.*; G. M. A. Grube, *Plato’s Thought* (Indianapolis, Ind.: Hackett, [1935] 1980), 138; R. C. Cross and A. D. Woozley, *Plato’s Republic: A Philosophical Commentary* (New York: St. Martin’s Press, 1964), 287. C. D. C. Reeve explicitly rejects this interpretation and yet invokes it at another point (C. D. C. Reeve, *Philosopher-Kings: The Argument of Plato’s Republic* [Princeton: Princeton University Press, 1988], 308 n. 35, 160).

14. The example comes from Adam’s commentary, *ad* 609e.

15. Chris Bobonich has suggested to me that besides the interpretation offered above, we could try to avoid the counterexamples by specifying natural evils at a much higher level of generality than Plato’s examples suggest. If wood’s natural evil is not just rot, but a broad category of things, then perhaps it is plausible to say that only the natural evil of wood can destroy wood. I believe that this strategy is overpriced, however. Not only do Plato’s examples of natural evils have to be read away, but also his argument loses much of its bite, for if each thing’s natural evil is construed very broadly, then there is little left to infer from the fact that any given thing’s natural evil cannot destroy it.

16. A general interpretation like (D) could be further specified parallel to (A)-(C) depending on whether the natural evil’s causal role in destroying *y* were taken to be solitary, primary, or merely partial. My wordings for (D) are intended to reflect my belief that, as the argument develops, Plato needs (D) to say that the natural evil of *y* must be capable of playing at least a primary causal role in the destruction of *y*. Moreover, I take it that the reasoning I shall suggest for the essential destructibility claim sustains this strong version of (D).

17. Those who still want to stick Plato to (A), (B), or (C) have two re-

sources at their disposal. First, one might point to the apparent inference (*gar*, “for”—609b1) of the essential destructibility claim from the claim that neither a good thing nor a neutral thing can destroy. If this is a strict inference, then Plato must be assuming that there are no evils for a given thing except its natural evil. Hence, (A) would be true. (And presumably, we would attempt to construe natural evils as broad classes, for which strategy see note 15 above.) Against this, I contend that the essential destructibility is more sensibly inferred not from the claim about goods and neutrals *via* the *gar*, but from background considerations *via* the *ara* in 609a9, as I explain in the text, below. On my reading, the *gar* is merely introducing a helpful, partially explanatory fact without signaling a strict inference. See J. D. Denniston, *The Greek Particles*, 2d ed. (New York: Oxford University Press, 1950), 61–62. (I thank Chris Bobonich for putting this objection to me very persuasively.)

Second, one who wants to stick Plato to (A), (B), or (C) might point to his parallel arguments that food’s natural evil (whatever it is) does not destroy the body (609e1–610a4) and that the body’s natural evil (illness) does not destroy the soul (610a5 ff.). At first blush, these arguments seem to assume something like (B), for he seems to say that if another’s natural evil is to kill one, it must do so by bringing about one’s own natural evil. Indeed, the unfortunate example at 609e1–610a4 is most naturally read as committing Plato to (B). I shall argue below, however, that the passage at 610a5 ff. is best understood as an *ad hominem* argument whose success depends on nothing like (A), (B), or (C), and since the parallel passage at 609e1–610a4 is subordinate to the argument of 610a5 ff. (see especially 610a5), we should be hesitant to stick Plato to (B) and impugn the whole argument on its authority.

18. Annas, *Introduction*, 345. This book is extremely admirable for the clarity with which it makes criticisms. In contrast to the many critics of this argument (and others) who do not bother to make their complaints explicit, Annas stimulates a direct, productive engagement with the argument. For this reason, I refer to her critique often.

19. Paul Shorey saw the connection between the Book Ten immortality argument and the Book One function argument (352d8 ff.), but he took both arguments to be guilty of the same waffle, because the soul has two functions, one to provide life and the other in terms of the moral life. Paul Shorey, *The Unity of Plato’s Thought* (New York: Archon Books, [1903] 1968, 41). Shorey is certainly right that the soul is traditionally construed by Greeks as a life-principle, but it is not clear to me that Plato must attribute to the soul two functions.

20. Stewart, *Myths*, 73.

21. Adam, commentary *ad* 609d.

22. Adam himself thinks (*ad* 609d) that Plato’s solution to the problems of evidence lies in the sort of *a priori* considerations which he finds in the *Phaedo*: “It is in fact the conception of soul as the principle of life which explains (from the Platonic point of view) Glaucon’s emphatic [answer].” This explanation certainly shows how the soul is conceived of as immortal, but it does not work as a solution to the present argument’s problems. As a putative solution, it only begs the question, defining the soul as the sort of thing which must be immortal.

I take Plato to be arguing for this conclusion.

23. Technically, Plato need only assume that if the soul perishes, it perishes at some separation from the body or another, not that if the soul perishes, it perishes at its one and only separation from the body. The weaker assumption is compatible with Adam’s general suggestion that the soul perishes only after several lifetimes (although incompatible with the possibility that the soul perishes during the afterlife). Even on this weaker assumption, Plato might reasonably ask for evidence of vice leading to the soul’s destruction, as presumably some embodied souls are not in their first but in their *n*th and last incarnations. In other words, even on this weaker assumption, Plato will be able to say that there is no evidence of the only sort we could have to substantiate the suggestion that vice destroys the soul. (This point makes the longer argument offered above against Adam’s alternative necessary only if that alternative is taken to be that the soul eventually perishes during the afterlife.) I am grateful to Malcolm Schofield for discussion of this point.

24. Annas, *Introduction*, 345. One can also find this criticism in Reeve, *Philosopher-Kings*, 160, and in Waterfield (Plato, *Republic*, trans., with notes and an introduction, Robin Waterfield [Oxford: Oxford University Press, 1993], 450).

25. In the *Phaedo*, Plato does give some effort to rejecting the view that the soul is merely a condition of the body (91c–95a).

26. Richard Lewis Nettleship, *Lectures on the Republic of Plato* (New York: St. Martin’s Press, [1902] 1961), 356. This view can also be found in Auguste Diès’ introduction to the *Republic* in the Budé edition (Platon, *Oeuvres Complètes*, t. 6 [Paris, 1965], cxvii–cxix).

27. Annas, *Introduction*, 345. This criticism also appears in Shorey’s introduction to the *Republic* in the Loeb edition (Plato, *The Republic*, 2 vols. [Cambridge, Mass.: Harvard University Press, 1935], 2:lxvii), and it seems to have been a matter of dispute among nineteenth-century German scholars, judging from the note in Adam’s commentary *ad* 608d ff.

28. This, of course, is part of Plato’s criticism of writing. See *Phaedrus* 275d4–e5.

29. For the defense of poetry’s banishment, notice the way some prior assumptions about Forms and psychology are used. Plato tells us right from the start of that discussion, and right from the start of Book Ten, that the arguments aimed at insiders: note the repetition of the verb “we have been accustomed” (*eiōthamen*) at 596a6 and b6.

30. I am grateful to John Palmer for bringing the Proclus passage (*in Rep*, 91/20–92/3 Kroll) to my attention, and to David Sedley both for objecting to my original (Proclan) reconstruction and for pointing out the possible Melissian parallel.

31. Annas, *Introduction*, 346.

32. DK 30 B 7 (from Simplicius *in Phys* 111, 18), translation after Kirk, Raven, and Schofield (*The Presocratic Philosophers*, 2d ed. [Cambridge: Cambridge University Press, 1983]).

33. This chapter originally grew out of a presentation given to the Ancient

Philosophy seminar at the University of Cambridge in Lent Term 1996, and I am very grateful to the participants in that seminar for comments and criticisms, especially to Myles Burnyeat, Mary Hannah Jones, John Palmer, Kenneth Sayre, Malcolm Schofield, and David Sedley. Additionally, for their comments on an earlier draft, I thank Elizabeth Asmis, Chris Bobonich, Ian Mueller, Malcolm Schofield, and the Executive Editor of *Apeiron*.

14

## The Argument for Immortality in Plato's *Phaedrus*\*

Thomas M. Robinson

One of the most condensed and abstruse arguments for soul's immortality is to be found at *Phaedrus* 245C-246A. It is not even clear whether "soul" here is meant to refer to soul collectively<sup>1</sup> or to individual souls;<sup>2</sup> perhaps it refers to neither, but simply to "soul in all its forms."<sup>3</sup> Certainly the very general form of the argument which follows would lead one to the latter conclusion, though "soul" must still be understood as "rational soul" or perhaps rational *part* of soul if one is to abide by the limiting characteristics outlined in 245C 2-4: "our first step . . . is to discern the nature of soul, *divine and human*, its experiences and its activities." If this description has any meaning at all, the phrase  $\pi\alpha\sigma\alpha$   $\psi\omega\chi\eta$  can hardly refer to World Soul (as Posidonius seems to have thought<sup>4</sup>), nor can it include the souls of flies (as Harpocration seemed to think<sup>5</sup>). It must be the noetic soul, as Hermias saw,<sup>6</sup> the  $\theta\epsilon\iota\sigma\omega$   $\tau\iota$  common to gods and men.<sup>7</sup>

Until recent times this argument for soul's immortality seemed to stem from an assertion that it was "eternally moving" ( $\alpha\acute{\epsilon}\iota\kappa\iota\eta\tau\circ\varsigma$ ).<sup>8</sup> The discovery of an Oxyrhynchus papyrus, however, which read "self-moving" ( $\alpha\acute{u}\tau\kappa\iota\eta\tau\circ\varsigma$ )<sup>9</sup> convinced Robin that this reading made much more sense of the argument, and he duly incorporated it into the Budé text.<sup>10</sup> The move has found a number of champions, notably Vollgraff, Bignone, Pasquali, Müller, Ross, and more recently J. L. Ackrill.<sup>11</sup> It is certainly true that with this reading one can reduce a complicated argument to the neater lines of an Aristotelian first-figure syllogism, but this

in itself ought perhaps to be grounds for suspicion. As Diana has pointed out, *αὐτοκίνητος* is a word found nowhere else in Plato;<sup>12</sup> the first incontrovertible instance of its usage is at Aristotle, *Phys.* VIII.5, 258a2. Plato tends rather to use phrases like “that which moves itself” (*τὸ αὐτὸ ἔκειται κινοῦν*). So it is at least possible that *αὐτοκίνητος* is the gloss of some commentator interested in reducing Plato’s looser arguments to a more terse and respectable logical format. Be this as it may, the original reading *ἀεικίνητος* ought, I think, to be defended. The diffuse series of arguments will then run as follows:

*Introduction* (=the argument in a nutshell)

All soul is immortal, because it is *ἀεικίνητος*, and anything *ἀεικίνητον* is immortal (245C5).

The meaning and implications of the word *ἀεικίνητος* are then brought out in two arguments:

1. (i) To call anything *ἀεικίνητον* is to call it a *αύτὸ κινοῦν*.  
 (ii) For if it were *not* a *αύτὸ κινοῦν* there would be a cessation of life and movement, and conversely, if it *is* a *αύτὸ κινοῦν* there can *never* be any cessation from movement, since a *αύτὸ κινοῦν* never “breaks contact” with itself.  
 (iii) Unstated conclusion: *αύτὸ κινοῦν* and *ἀεικίνητον* are one, and presumably immortal.
2. (i) The *ἀεικίνητον* never had a beginning, for it is a “source of motion” (*ἀρχὴ κινήσεως*), and *ἀρχαῖ* cannot come to be from something else without ceasing to be *ἀρχαῖ*.  
 (ii) *Qua ἀρχὴ* it is also *ipso facto* indestructible, since “if a first principle were to be destroyed, nothing could come to be out of it, nor could anything bring the principle itself back into existence, since that first principle is needed for anything to come into being.”

A third argument links together the results of the first and second arguments, corroborating them with a *per impossibile* consideration for good measure. It runs as follows:

3. (i) *ἀρχὴ κινήσεως* and *αύτὸ κινοῦν* are one and the same (since both are identical with what is eternally in motion).  
 (ii) From this it follows that there will be no cessation from being

and no coming into being.

(iii) If this were not the case, the whole universe would sooner or later collapse into immobility.

These results are now applied to the notion of soul itself, and the fourth and final argument runs as follows:

4. (i) It has already been seen that the *αύτὸ κινοῦν* is immortal (i.e., by implication at the end of the first argument).  
 (ii) This will mean that soul, too, is immortal. For body, whose movement is from without, is as such “soulless” (*ἀψυχὸν*), whereas soul has as its nature movement which is from within.  
 (iii) One must therefore conclude that *αύτὸ κινοῦν* and soul are one and the same, and that soul will thereby be necessarily without beginning and immortal.

While Hackforth defends the original reading *ἀεικίνητος*, he can hardly be right in assuming that it is an *ἔνδοξον*<sup>13</sup> serving as a major premise, since the greater part of the argument is spent in elucidating it and outlining its implications. In Ackrill’s words, “it (i.e., the major premise) can perfectly well express a proposition required for the main proof though itself needing to be established by a subordinate proof.”<sup>14</sup> This “subordinate proof” in fact follows three stages:

1. The *ἀεικίνητον* is a *αύτὸ κινοῦν*.
2. A *αύτὸ κινοῦν* is immortal.
3. Soul is a *αύτὸ κινοῦν* and thereby immortal.

In the total argument as I have outlined it, the first two proofs are apodeictic, and each one sufficient of itself to show the immortality of *ἀεικίνητον* and *αύτὸ κινοῦν*. But Plato chooses to stress their cumulative rather than their individual plausibility, and adds what Hermias calls an argument *per impossibile* for good measure.<sup>15</sup>

The argument, as many have pointed out, has much in common with the final argument for immortality in the *Phaedo*.<sup>16</sup> But whereas that of the *Phaedo* sprang from a metaphysic of Plato’s own creation, the *Phaedrus* argument is of a more empirical cast, and has its roots deep in the pre-Socratic tradition.<sup>17</sup> It also goes a long way towards solving difficulties presented by the psychology of the *Phaedo*. There the soul had been assimilated to the Ideas, and the hiatus between the unmoving intelligible

world and that of sensible things subject to motion had meant an inadequate explanation of the soul as we experience it in the world of movement. In the *Phaedo* he had found himself compelled to run in the face of all philosophic tradition, as part of his defense of his newborn metaphysic; the elementary attribute of all living things, movement, was quietly shelved, and the static, homogeneous, unchanging entity whose immortality he was left to prove was recognizable only to himself. Now soul is seen as the source of motion or activity, and a more balanced appreciation of its true nature is possible. This is not to argue, of course, that the notion of movement appears like a bolt from the blue in the *Phaedrus*. It was there by implication in both *Phaedo* and *Republic* when soul was seen as a principle of Life, or bound up with the Idea of Life;<sup>18</sup> unfortunately this notion was allowed to stand on equal terms with others barely compatible with it—such as that of soul's likeness to the static Ideas<sup>19</sup>—and one is led to conclude that Plato was not then fully aware of the implications of what he was asserting.

A much-discussed question is the relationship of this doctrine of soul as source of motion to the problem of the precosmic chaos at *Tim.* 49A ff., and a number of scholars in recent years have used the one to explain the other.<sup>20</sup> For at 53A2 in particular we read how the Receptacle, itself in motion, moves the four “kinds” ( $\gammaένη$ ) which it has received into itself, thus separating the “unlike” sorts farthest from one another, and bringing “like” kinds close together; so that even in this precosmic world there is *some* sort of organization, some region to which each “kind” gravitates. Before that they are “without plan” (or “without pattern,” or “without proportion,”  $\delta\lambda\gammaως$ <sup>21</sup>) and “without measure” ( $\delta\mu\acute{e}τρως$ <sup>22</sup>). It is the action of the Receptacle itself, apparently, which gives them some sort of primitive organization. It is when they are in this latter state—a state qualified as “what one would expect from a thing when  $\vartheta\epsilon\circ\varsigma$  is absent from it”<sup>23</sup>—that the Demiurge intervenes. The point at issue is whether the power of motion here expressly attributed to the Receptacle indicates the presence of soul or not. One can begin by saying that in the *Phaedrus* and *Laws* the argumentation is very generalized and schematic: no reference is made to precosmic states, and understandably so, since the context does not demand them.<sup>24</sup> In addition, the reader coming new to the text can be forgiven for assuming that in these two dialogues the movement referred to is that which he sees operating in the world around him, and that the “bodies” in question are those with which he is acquainted in the world of everyday experience: in a word, the movement which is to be found in an (already) organized universe. How far it would

apply, if at all, to a precosmos is hard to see. It was argued by Plutarch that it definitely does apply to the precosmos;<sup>25</sup> by most modern scholars, following Proclus, that it applies to that element in the cosmos which talk of a precosmos is meant to symbolize;<sup>26</sup> by Herter that it ought logically to have applied to the precosmos (real or metaphorical) but that it does not appear to do so.<sup>27</sup> Certainly the passage 52E-53C (like 30A3-5) offers no hint of the presence of soul, rational or irrational. While it must be admitted that the *argumentum e silentio* is often a foundation on quicksands, in this particular instance it seems to have more cogency than usual. If the precosmic chaos did possess soul, that soul is part of what the Demiurge “took over”; he did not create it. It will also be an irrational soul, since it does not possess the exclusively rational circular motions with which (rational) World Soul is endowed.<sup>28</sup> These conclusions combined lead one to doubt any thesis that such an irrational soul is to be included within World Soul as outlined in the *Timaeus*, since it is apparently both distinct from it and different in kind from it.<sup>29</sup> If, however, one still accepts the notion of soul as the motive force of the precosmic chaos, though not as “part” of World Soul, one is left with the conclusion that there are or were two distinct psychic forces in the universe, one rational, one irrational. But in the *Politicus* myth Plato rejects the notion of two opposing divinities in the universe,<sup>30</sup> and one could argue that the notion of two opposing “psychic agents” would be open to the same objection. (On the other hand, there is no particular reason why Plato should not have changed his mind; the problem is sufficiently puzzling to give pause to the most confident.) The most basic objection to the theory, however, seems to be as follows: the notion of a soul other than that created by the Demiurge is so crucial in its implications that it seems incredible that Plato should not have given clear indication of its presence. To say that one can “infer” its presence from the evidence of the *Phaedrus* and *Laws* is not enough. For, if Owen and others are right, both *Phaedrus* and *Laws* may well have been written after the *Timaeus*,<sup>31</sup> and if this is the case, one would be engaged in the dubious task of interpreting an earlier dialogue in the light of later ones. But even if, for the sake of argument, one granted that the *Phaedrus* predates the *Timaeus*, it is still far from evident that both are talking about the same thing. In the *Phaedrus*, for example, it is natural to assume that the movement in question is that which obtains among bodies in the fairly organized cosmos known to us by sense experience; to infer that it also applies to the pre-cosmic world where duration is not Time and the bodily (i.e., organized bodies which we can recognize) is as yet nonexistent, is less easy.

The *Phaedrus* seems to be saying that soul is the cause of all movement in an organized world, a world measurable by Time. In a nonorganized world *not* measurable by Time one can wonder whether the movement in question has anything to do with this.<sup>32</sup> At this stage words start to break under the strain. Plato is compelled to give *some* description of the precosmic chaos, and talk of movement in such a world is no more and no less intelligible than phrases like "before this" ( $\piρὸ τούτου$ ) (53A8) in the same passage, when Time has been admitted to be absent. So Herter seems right in saying that the doctrine of *Phaedrus* and *Laws* is not to be applied to the *Timaeus*, but perhaps unjust in saying that logically it ought to have been. Plato, as far as one can see, is dealing in two instances with two completely different types of motion, the one accepted and universally admitted, and operating in an organized world of temporal succession, the other a *pis aller* trying to describe a reality in every sense different. If this analysis is correct, there seems no reason to reproach Plato here with lack of logic, though on points of clarity he may leave a lot to be desired.

## Notes

\* This is a revised version of a paper presented at the 1966 meeting of the Society for Ancient Greek Philosophy in Toledo, Ohio. A modified version has appeared in *Apeiron* II (July, 1968): 12-18. By permission of the editors.

1. A.-J. Festugière, "Platon et l'Orient," *Revue de philologie, de littérature, et d'histoire anciennes* 3 (1947): 21. Like Posidonius, he takes the argument to refer to World Soul.

2. P. Frutiger, *Les Mythes de Platon* (Paris: Alcan 1930), 130-34.

3. J. B. Skemp, *The Theory of Motion in Plato's Later Dialogues* (Cambridge: Cambridge University Press, 1942), 3 n. 1, and A.-J. Festugière, *Revue des études grecques* 59/60 (1946-47): 496.

4. Hermias, *Commentary*, ed. P. Couvreur (Paris: É. Bouillon, 1901), *ad loc.*

5. Hermias, *Commentary*.

6. Hermias, *Commentary*.

7. *Timaeus* 41C 7, 69C 3.

8. 245C 5.

9. *Ox. Pap.* 1016.

10. See n. 11.

11. For the references and comment (both on these and on L. Robin) see C.

Diano, "Quod semper movetur aeternum est," *La Parola del Passato* 2 (1947): 189-92. For J. L. Ackrill's comment see review of R. Hackforth, *Plato's Phaedrus* (Cambridge: Cambridge University Press, 1952), *Mind* 62 (1953): 278.

12. Diano, "Quod semper," 190 ff. For a number of recent comments on the subject see H. Cherniss, "Plato (1950-57)," *Lustrum*, 4-5 (1959), Band 4, 137. The most telling criticism of Robin's arguments for  $\alphaὐτοκίνητος$  is that of Diano, followed by W. J. Verdenius, "Notes on Plato's *Phaedrus*," *Mnemosyne* 4 (1955): 267. Robin has made the opening lines into an Aristotelian syllogism, taking  $\delta\epsilon$  at C6 as adversative, and  $\delta\eta$  at C7 as consecutive (see his translation *ad loc.*, Budé edition). Diano points out how  $\delta\epsilon$  can equally well be progressive, and  $\delta\eta$  emphatic, thereby changing the character of the argument considerably. At C5 the words  $\tauὸ γαρ \alphaεικίνητον \deltaθάνατον$  are really a definition, as Diano sees (191). As such, it is the argument's conclusion, rather than its starting-point. It gains by being placed startlingly at the beginning, in the manner of a textbook definition; what follows is an *explicatio* and *probatio* of what it contained succinctly.

13. R. Hackforth, trans., *Plato's Phaedrus* (Cambridge: Cambridge University Press, 1952), 65.

14. Ackrill, review of Hackforth, 278.

15. Hermias, *Commentary*, 103.3 ff.

16. On the affinities between the two arguments see Hackforth, *Plato's Phaedrus*, 68, who refers to Skemp, Frutiger, and Bury. Add H. Cherniss, *Aristotle's Criticism of Plato and the Academy*, I (Baltimore, Md.: Johns Hopkins Press, 1944), 435-38, and H. Cherniss, "The Sources of Evil According to Plato," *Proceedings of the American Philosophical Society*, 98 (1954): 28. These scholars take it for granted that Plato got his idea of psychic movement from the observation of living things (e.g., Hackforth, *Plato's Phaedrus*, 68, and Cherniss, *Aristotle's Criticism*, 435-38). Festugière, however, has argued that he took it from "la découverte purement grecque de la régularité des mouvements célestes" ("Platon et l'Orient," 21). In this he is followed by V. Martin, "Sur la condamnation des athées par Platon au Xe Livre des *Lois*," *Studia Philosophica* 11 (1951): 116, who also allows some cogency to the accepted explanation (117). If this guess is correct, the line of reasoning seems to be as follows:

- a. the circular motion of the heavenly bodies manifests the operation of intelligence ( $νοῦς$ ) (*Timaeus* passim).
- b. intelligence cannot arise apart from soul (*Timaeus* 30B3; *Sophist* 249 A4-8).
- c. soul, with intelligence, is consequently responsible for the movement of heavenly bodies.
- d. by analogy, the same reasoning will be applicable to *all* noetic souls, celestial or otherwise.

The strength of this argument, if it genuinely represents Plato's thought, is that it accounts for the motive power of the *noetic* soul (see nn. 6 and 7 of this pa-

per)—i.e., the *only* soul apparently in question in the *Phaedrus* argument. In the *Laws*, however, the soul which is the source of movement is something much wider than the noetic soul: it is in itself neutral, only taking on ethical color when seen as νοῦν προσλαβόντα or ἀνοίᾳ συγγενομένη. This seems to lay stress on the nature of soul as a vital principle, the *sine qua non* of all activity, good or bad (V. Martin, "Sur la condamnation," 120), rather than soul as essentially characterized by intelligence and "care" (ἐπιμέλεια) (the *Phaedrus* position), and here Festugière's explanation hardly fits.

17. For a study of the doctrine of movement in Empedocles, Alcmaeon, and Pythagoreans see Skemp, *The Theory of Motion*. (His views on Alcmaeon are criticized by Festugière, "Les 'Mémoires Pythagoriques' cités par Alexandre Polyhistor," *Revue des études grecques* 58 [1945]: 59-65.) H. C. Baldry ("Embryological Analogies in pre-Socratic Cosmogony," *Classical Quarterly* 26 [1932]: 27-34) has shown, I think convincingly, that Anaximander saw the world as a living creature, in many ways like a fetus. For an account of soul as source of motion in Alcmaeon, Anaxagoras, and Diogenes of Apollonia, see A. Diès, *Autour de Platon II* (Paris: Gabriel Beauchesne, 1927), 536. For Aristotle's views on soul as a principle of motion among pre-Socratic thinkers, see *De An.* 403b24 ff., and compare *Phys.* VIII.9, 264b17-266a5.

18. Cherniss seems to me to go beyond the evidence when, in seeing a connection between the final proof of the *Phaedo* and that in the *Phaedrus*, he identifies the "idea of life" with the "idea of motion" (*Aristotle's Criticism*, 435-38). Compare H. Cherniss, review of A.-J. Festugière, *Gnomon* 22 (1955): 208, where he makes the point that the notion of self-motion is to be found as early as *Charmides* 168E. To have a notion, however, is not necessarily to see its implications. If, at the time when he wrote the *Phaedo*, Plato really did see all the implications of the notion of the "Idea of Life," it is hard to think that he would have painted the very static portrait of ψυχή that he did. For a different viewpoint see G. Müller, *Studien zu den platonischen Nomoi* (München: Beck, 1951), 79-80.

19. *Phaedo* 80B1 ff.

20. Skemp, *The Theory of Motion*, 77-78, and Cherniss, "The Sources of Evil," 25. Both use the same notions of "first order" (πρωτουργός) and "second order" (δευτερουργός) movement at *Laws* 897A4-5 to prove rather different things. For Skemp, the "second order" movements are the "necessary conditions" (συναίτια) of the *Timaeus* and *Phaedo*; that is, they are the purely mechanical *sine quibus non* of the (psychic) actions of Intelligence and Necessity. In this way the *Laws* is used to bolster a view of soul, rational or irrational, as the direct (πρωτουργός) cause of all movement. For Cherniss, (rational) soul is the *indirect* (though ultimate) cause of all movement: this is based on his assumption that the "first order" and "second order" movements of *Laws* 897A4-5 stand for "primary" and "secondary" causation respectively, and that they find their exact parallel at *Tim.* 46E in the causes:

- (a) "which work with intelligence to produce what is good and desirable" and
- (b) "which, being destitute of reason, produce their sundry effects at random and without order."

On his analysis, (b) are the side-effects of (a), and, while serving no directly rational end, find their *fons et origo* in rational soul all the same. The point to notice is that soul is taken to be rational; Cherniss will not concede that there is an irrational World Soul (Skemp) or part of World Soul (F. M. Cornford, *Plato's Cosmology* [London: Routledge & Kegan-Paul, 1937], 208) represented by Necessity (Skemp) or by the Different (Cornford). See also H. Herter, "Bewegung der Materie bei Platon," *Rheinisches Museum für Philologie* N. F. 100 (1957): 332.

- 21. 53A8.
- 22. 53A8.
- 23. 53B3-4.
- 24. See G. Vlastos, "The Disorderly Motion in the *Timaeus*," *Classical Quarterly* 33 (1939): 78.
- 25. *De. an. procreat. in Tim.* 1016C-D.
- 26. Proclus *In Tim.* 323B (Diehl III, 273/31); J. Moreau, *L'Âme du Monde de Platon aux Stoïciens* (Paris: Société d'édition "Les Belles lettres," 1939), 25 ff., 39 ff.; W. F. R. Hardie, *A Study in Plato* (Oxford: Clarendon, 1936), 152; A. E. Taylor, *A Commentary on Plato's Timaeus* (Oxford: Clarendon, 1928), 115 ff., 390 f.; F. M. Cornford, *Plato's Cosmology*, 207.
- 27. Herter, "Bewegung der Materie bei Platon," 343 n. 55, and 346.
- 28. Several scholars have taken it for granted that the Circle of the Different in World Soul is or represents irrationality. See L. Robin, *Platon* (Paris: Presses universitaires de France, 1935), 328; F. M. Cornford, *Plato's Cosmology*, 76, 96, 118-19, 144 n2, and 208; A. D. Winspear, *The Genesis of Plato's Thought* (New York: S. A. Russell, 1956), 332; and G. Morrow, "Necessity and Persuasion in Plato's *Timaeus*," *Philosophical Review* 59 (1950): 162-63. But there is no indication in the text that the composition of the one Circle is in any way different from that of the other. True, both have different provinces, different spheres of influence. The Circle of the Same has to do with the eternally stable world of the Ideas, a world only penetrated by pure intelligence, while that of the Different deals with the subrational world of sense-experience. But it still approaches the world of sense in the most rational manner compatible with that exiguous amount of reality which such a world enjoys, and compatible with such a world's inherent inability to be fully comprehended by rational analysis. Given the nature of the subject-matter, the Circle of the Different will never come to conclusions other than δόξαι and πίστεις, but its *rational* activity is made clear by the way it can guarantee that any such δόξαι and πίστεις will be factually correct. See J. B. Skemp, *The Theory of Motion*, 208-209 n. 5, and J. Gould, *The Development of Plato's Ethics* (Cambridge: Cambridge University

Press, 1955), 199.

29. No doubt this is one reason why Aristotle (*Metaph.* Λ 6, 1072a1-2) says that Plato's αὐτὸν ἔαυτὸν κινοῦν is only ἐνίοτε ἀρχήν; as he goes on to say, ὅτερον γὰρ καὶ ἄμα τῷ οὐρανῷ ἡ ψυχή, ὡς φησίν. See Herter, "Bewegung der Materie bei Platon," 329 n6, and A. Rivaud, *Le problème du devenir et la notion de la matière dans la philosophie grecque depuis les origines jusqu'à Théophraste* (Paris, 1906), 337. T. Gould (*Platonic Love* [London: Routledge & Kegan-Paul, 1963], 29) suggests that the precosmos is described as in chaotic motion to make the most complete contrast with the static permanence of the Ideas, where Intelligence is supremely manifest.

30. 270A1-2.

31. For the arguments, see G. E. L. Owen, "The Place of the *Timaeus* in Plato's Dialogues," *Classical Quarterly* N.S. 3 (1953): 79-95. He is followed by J. Gould, *The Development of Plato's Ethics*, 202 n. 3; D. W. Hamlyn, "The Communion of Forms and the Development of Plato's Logic," *Philosophical Quarterly* 5 (1955): 290 n. 3; D. A. Rees, "Bipartition of the Soul in the Early Academy," *Journal of Hellenic Studies* 77, Part I (1957): 113 n. 29; and C. Strang, "Plato and the Third Man," *Proceedings of the Aristotelian Society* Supp. 37 (1963): 147-64. W. G. Runciman (*Plato's Later Epistemology* [Cambridge: Cambridge University Press, 1962], 4), while still somewhat diffident, feels that Owen has at any rate tested the accepted dating of the *Timaeus*. For extended criticisms of Owen's view, see H. Cherniss, "The Relationship of the *Timaeus* to Plato's Later Dialogues," *American Journal of Philology* 78 (1957): 225-66, and J. Rist, "The Order of the Later Dialogues of Plato," *Phoenix* 14 (1960), 207-21. See also the postscript in J. B. Skemp's edition of the *Politicus*. For a startling piece of statistical evidence in favor of Owen's thesis, see D. R. Cox and L. Brandwood, "On a Discriminatory Problem Connected with the Works of Plato," *Journal of the Royal Statistical Society Series B*, 21 (1959): 195-200.

32. Cornford (*Plato's Cosmology*, 205) finds "bodily changes" in the precosmic chaos, and so argues (from the *Phaedrus* and *Laws* X) to the presence in it of an (irrational) soul. In this he seems to stress the wrong alternatives in the two ambiguous terms "change" and "bodily" which Plato is using. It is true that the Demiurge is said to "take over" all that is *visible* (30A3-4), but one need hardly conclude, with Proclus, that the bodily must therefore be present in the precosmic chaos, if only in some minimal way. (See Cornford, *Plato's Cosmology*, 205.) Plato himself only talks about ἥχνη and δυνάμεις, and there is no reason to think they necessarily obey the same rules as organized bodies (52E-53B). The word 'visible' is probably used loosely for "nonspiritual" or "nonpsychic," in much the same way as in the *Phaedo* soul is called ἀειδές.

## Postscript 2000

For an important article on the *Phaedrus* argument for immortality that takes issue with parts of my own interpretation of it readers should consult Richard Bett (this volume, 335-62). His interpretation of soul in the argument as 'stuff' seems to me attractive, not least because when Plato returns to the topic (in *Laws* 10) he quite possibly had just such a view of soul in mind. But it seems to me less clear that the argument is meant to cover more than just *rational* soul-stuff, despite the fact that the myth of the dialogue (like all the eschatological myths) gives the strong impression that the soul which survives the body, and for that matter the souls of the gods themselves, is a soul complete with a full emotional apparatus.

My own instinct is still to see the immortality argument of the *Phaedrus* as describing a soul understood as *dianoia* only (see *Phdr.* 239a5, 239c1, 249c5). If, as might well have happened, friends and pupils pointed out to Plato that this was very difficult to reconcile with a natural understanding of the image of soul portrayed in the myth of the dialogue (and other eschatological myths), this would offer a plausible reason for features of his subsequent argument for immortality in the *Laws*, which could reasonably be interpreted as evidence of an attempt by Plato to rethink the matter in a way that would accommodate the difficulty.

## Immortality and the Nature of the Soul in the *Phaedrus*

Richard Bett

“First we must inquire what kind of soul the discussion is about.” So begins Hermeias’ commentary on the argument for immortality in the *Phaedrus* (245c5-246a2).<sup>1</sup> It is a piece of advice that has not, I think, been sufficiently heeded by modern scholars. There has been some discussion of various textual problems and points of translation, and this has, of course, involved some scrutiny of the argument’s structure and presuppositions; but I am not aware of any really detailed, step-by-step analysis of the argument—other than Hermeias’ own.<sup>2</sup> As a result, the precise relation between Plato’s views on immortality and the nature of soul in the *Phaedrus* and in other dialogues has not received as close attention as it could. In particular, the question of what conception of soul Plato is operating with in this argument has tended to be dealt with too superficially—by examination of single, troublesome phrases, rather than by assessment of the argument as a whole.

My purpose, then, is to analyze the *Phaedrus*’ argument for immortality as minutely as I can. This argument is of course of considerable intrinsic interest, and I shall begin by treating it largely as an isolated piece of logic, examining the nature and cogency of the various premises and inferences. However, the argument must also, of course, be seen in its context. By examining it in detail, I hope also to clarify, to some degree, the place of the *Phaedrus* among Plato’s works—or at least, among those works that deal with questions about the soul and immortality. In

addition, I hope to arrive at a better understanding of the argument's place in the *Phaedrus* itself. For another neglected issue, I believe, is that of the relation between the argument for immortality and the myth of the charioteers that immediately follows it. Typically, the two sections have been discussed too much in isolation from one another; yet clearly the nature of the soul is a central topic for both of them.

That Plato intends the argument as a rigorous proof, and hence as admitting of the detailed examination I intend to give it, is not, I think, in doubt. In this respect it contrasts sharply with the myth that follows, and Plato purposely signals the contrast in two ways.<sup>3</sup> First, there is the warning at 246a6 (immediately preceding the myth) that what follows is a description of the soul only ὡς ἔστι, not οἷον ἔστι; the proof has no such qualifications attached to it. Secondly, the styles of the two passages are very different. The proof is presented in extremely spare, choppy prose; Plato here seems to be aiming for maximum clarity and logical perspicuity.<sup>4</sup> When he embarks on the myth, his language loosens up, and seems positively florid compared with the rigor and economy of the previous passage. In fact, of all the arguments for immortality in Plato, this one is much the closest to what we would consider a formal proof. So in dissecting it, we are not approaching it in a spirit any different from that in which Plato composed it.

## I

The basic shape of the argument is as follows:

- (1) Soul is that which is its own source of motion.
- (2) That which is its own source of motion is immortal. Therefore
- (3) Soul is immortal.

Plato presents it, though, in the reverse order. The conclusion (3) is stated at the beginning (245c5); the argument for (2) occupies him from there until e2, and (2) is stated at e2-3; (1) is then argued for between 245e3 and 246a1; and finally, at 246a1-2, the conclusion of the whole argument is restated. The argument for (1) is relatively straightforward, and we shall deal with it later; (2), however, is established in a much more complex fashion, by means of two independent subarguments. The premises of the first subargument (which I shall label A) are as follows:

- (Ai) That which is its own source of motion is *always* in motion.
- (Aii) That which is always in motion is immortal.

The premises of the second subargument B are:

- (Bi) That which is its own source of motion is a source of motion for everything else that moves.
- (Bii) That which is a source of motion for everything else is ungenerated and imperishable.

A is also presented in the reverse order, but B occurs in the order in which I have just exhibited it.<sup>5</sup> I shall now analyze the steps one by one, in the order in which Plato gives them, beginning with subargument A.

We begin with a statement of (Aii)—τὸ γὰρ ἀειχίνητον αθάνατον (245e5). No argument is offered for this premise; it seems to be simply taken for granted.<sup>6</sup> Presumably the underlying thought is that something which is always in motion must always exist, and so must be immortal. Provided “always in motion” means “in motion for eternity” (on which more below), this seems hard to quarrel with. It is only when we move to (Ai) that the real argumentation begins.

The effect of (Ai) is to identify *that which is always in motion* with *that which moves itself*; given Aii), this establishes the main premise 2). Plato now states (Ai) at 245c7-8—“only that which moves itself . . . never ceases moving”—prefacing it with what amounts to a statement of its converse—“that which . . . is moved by something else, since it admits of a cessation of movement, admits of a cessation of life” (c5-7).<sup>7</sup> At the same time, he offers a reason for the assertion; this is contained in the terse and cryptic phrase ἀτε οὐκ ἀπολεῖπον ἔαυτό (c7-8).

What is meant by ἀτε οὐκ ἀπολεῖπον ἔαυτό? Literally, the phrase must mean something like “inasmuch as it does not abandon itself”; and the natural way to understand this is as saying that for a self-mover to cease moving would be for it to abandon its own nature. It is, then, an essential property of a self-moving thing (to speak in a later, but surely an appropriate, terminology) that it be moving itself *at any given time*; self-moving things are necessarily in *constant* self-motion. Two questions now arise. First, why does Plato think that this is so? And second, if we accept that it is so, does this suffice to show that self-moving things are *eternally* in motion? On the first point, Plato offers us no further enlightenment; but perhaps the underlying idea is something like the following. Suppose that a self-mover were to cease to move itself. Then

it could not start moving itself again unless galvanized into motion either by itself or by something else. But if it were galvanized into motion by itself, this would have to occur by the agency of some part of itself that was already in motion—in which case, contrary to hypothesis, the thing would not really have ceased to move itself after all. But if by something else, it could no longer count as a *self*-mover. Thus something which truly moves itself must move itself continually.<sup>8</sup>

Suppose that we accept this conclusion (whether or not Plato arrived at it in the way I just proposed). The other question is whether this is equivalent to saying that something which moves itself does so eternally; and to this the answer is clearly “no.” As we saw, a straightforward understanding of premise (Aii) requires that we take *ἀεικίνητον* in c5 as meaning “eternally in motion”; so if the argument is to work, *οὐποτε λήγει κινούμενον* in c8 must presumably have the same sense. However, it is not in *this* sense that Plato has shown us that self-movers “never cease moving.” If being in constant self-motion is an essential property of a self-mover, then a self-mover cannot cease to be in motion *and still be a self-mover*; but this is not to say that it will necessarily keep moving *for ever* (and so existing for ever).

There is, then, a gap in the argument. At the same time, however, it is quite understandable that Plato should not have been aware of it. The trouble is that “X never ceases to be F” (of which “that which moves itself never ceases moving” is an instance) can mean *either* “It is impossible that there should be an X which is not F” (i.e., F is an essential property of Xs) *or* “Xs are eternally F”; the second entails that Xs are immortal, but the first does not. It is only the first that Plato is really entitled to in this instance; for as I understand it, *ἄτε οὐκ ἀπολεῖπον ἔαυτό* makes a point about an essential property, not about the eternal possession of any property. But it is not surprising that the ambiguity should have escaped him.

It is interesting that precisely the same illegitimate move is made in the final argument for immortality in the *Phaedo*—an argument with which this one is often compared; indeed, I take this parallel as confirming evidence for my reading of what is going on here in the *Phaedrus*.<sup>9</sup> In the *Phaedo* Plato moves from “The soul always comes bringing life to whatever it occupies” (105d3-4), and its corollary (d10-11) “The soul may never admit the opposite of that which it always brings with it” (i.e., death, as he goes on to say) to “The soul is immortal” (e6). Again, it is in one sense uncontroversial, given the commonsense Greek view of the soul, that “the soul always brings life.” That is, anything which has a soul

is thereby necessarily alive; it is an essential property of souls that they are not present in organisms which have died. But this is not to say that souls exist eternally. Nothing can cease to bring life *and still be a soul*; this follows from the essential property of souls we have just noticed, and in this sense we may agree that the soul does not “admit death.” It does *not* follow that the soul must continue to possess this life-giving capacity throughout time, and so be immortal. As in the *Phaedrus*, the difficulty stems from Plato’s use of the words *ἄτε* and *οὐποτε*—words whose primary senses have to do with temporal duration—to talk about what we would call essential properties.<sup>10</sup> It is hard to know how else he could have expressed himself, given the vocabulary available to him; but the resulting ambiguity was a treacherous one.<sup>11</sup>

So much, then, for the argument given in the phrase *ἄτε οὐκ ἀπολεῖπον ἔαυτό*. While the argument is not one we can accept, it is significant for the picture of the soul that it obliges Plato to hold. He must embrace a view of the soul as not simply a self-moving thing, but as *necessarily* in *unceasing* self-motion. We will return to this point later; for the moment, let us continue with the analysis of the argument. Subargument A for the main premise (2) is now complete, and Plato moves on to subargument B; the discussion here is somewhat more lengthy.

Premise (Bi) is stated in the same sentence we have been looking at for some time; “only that which moves itself . . . is a source and first principle of motion for the other things that move” (c7-9).<sup>12</sup> (Bi) is not argued for directly. Instead, Plato immediately states one half of (Bii)—*ἀρχὴ δὲ ἀγένητον* (d1)—and the argument for (Bii) takes him from there to 245e2. (Bii) is stated in full at d7-8; “it is not possible that this [i.e., that which moves itself] should either perish or come to be.” However, the first point he presents in favor of (Bii)—namely, that everything which comes to be must do so from a first principle which does not itself come to be from anything else (d1-2, repeated at d6)—is also a point in favor of (Bi) (provided we understand *γένεσις* as a species of *κίνησις*—I shall return to this later). Indeed, as if to signal that he has been confirming (Bi) at the same time as he has been arguing for (Bii), Plato restates (Bi) at d6-7; “thus that which moves itself by itself is a first principle of motion.”

The argument for (Bii) is as follows. (I have just mentioned the first stage of it, but I shall repeat this for clarity’s sake.) A first principle is ungenerated (*ἀγένητον*), since everything that comes to be comes to be out of a first principle, and it (i.e., the first principle) does not come to be out of anything else (d1-2); if a first principle did come to be out of

something else, then “it would no longer be out of a first principle that all that comes to be comes to be” (d2-3).<sup>13</sup> This last point is somewhat obscure, and the text may be faulty. But the general idea, that a first principle is necessarily not generated from something else, seems clear enough; and it does follow plausibly enough that a first principle is ungenerated, assuming that nothing can generate itself. Hence, Plato continues, it must also be imperishable (d3-4); for (in view of what has just been said) if it were to perish, nothing else could bring it back into being, nor could anything else come to be out of it (d4-6). One might still ask why this shows that it could not perish. Plato gives his answer at d8-e2, in conjunction with his explicit statement of the premise (Bii) with which we are now concerned. Essentially, the point is that if it were to perish (so bringing to an end all possibility of γένεσις), the universe would collapse into immobility.<sup>14</sup> However, it is not stated why this, in turn, could not happen.<sup>15</sup>

This concludes subargument B for the main premise (2); immediately after, at e2-3, (2) is explicitly stated. What are we to make of subargument B? There are two points about it which I find of interest. First, the argument requires that we think of γένεσις as a species of κίνησις. For Plato introduces the notion of an ἀρχή κινήσεως, and then immediately goes on to say (by way of arguing for (Bii)) that “everything that comes to be comes to be” out of an ἀρχή. For the argument to work, it must obviously be the same kind of ἀρχή under discussion throughout; and it follows that κίνησις cannot simply mean “(loco)motion,” but must refer to any kind of change whatever. However, there is nothing very startling in this. We find just such a broad notion of κίνησις explicit in the *Laws* (again in the context of a discussion of soul as self-mover); at 893b6-894c8 ten kinds of “motion” (κίνησις, also interchangeably μεταβολή) are distinguished, two of which are γένεσις and φθορά (see, e.g., 894b1).<sup>16</sup> Similarly, in *Physics* III.1 Aristotle gives his definition of motion (κίνησις) as the fulfilment of a potentiality qua potentiality, and immediately goes on to distinguish various species of motion, among which are γένεσις and φθορά (201a10-16). And indeed, a strong connection between γένεσις and κίνησις occurs as early as Parmenides: “That which is” in Parmenides is explicitly said to be both ἀγένητον and ἀκίνητον, and these two properties appear to be seen as closely connected—see especially DK B8.26-8. Parmenides does not appear to treat γένεσις as actually a species of κίνησις; but in his use of these notions, he certainly seems to foreshadow those who do. There are,

then, respectable parallels for the idea that “coming to be” is a kind of κίνησις.<sup>17</sup>

On the other hand, this does not mean we should simply grant Plato the idea and move on. It may be understandable that he should have regarded γένεσις as a species of κίνησις; but I think that questions can still be raised about the role this assimilation plays in the argument. Recall that Plato has just claimed that a self-mover is a first principle of motion for everything else that moves. We have now discovered that “motion” includes “coming to be”; but presumably (though this is not explicitly stated here) it includes other things as well—at least some of the other kinds of “motion” mentioned in the *Laws* and in Aristotle’s *Physics*. But if this is so—if there are several different kinds of κίνησις—one might wonder why the same thing should have to be a first principle for all of them. If we accept that there are such things as self-movers, we can also accept that these are responsible for the “motion” of everything else. But it is not clear why any *one* self-mover should have to be responsible for every one of the different species of κίνησις. For example, why should not one thing (or kind of thing) be a first principle of γένεσις and some other thing (or kind of thing) a first principle of locomotion? Plato’s argument requires that anything which is a self-mover is a first principle specifically of γένεσις; but it is not obvious to me why this should be taken for granted. Needless to say, Plato does not respond to this query; and in view of the parallels I just adduced from Parmenides, Aristotle, and elsewhere in Plato, this should no doubt not surprise us. Among other things, these parallels demonstrate a pervasive tendency in Greek thought to see all the processes subsumed under the heading of κίνησις, in the broad sense we have examined, as being very intimately related. (In Parmenides and Plato, at least, this tendency results from the central contrast in both their philosophies between the world of change and the world of the changeless, the world of τὸ ὅν; someone for whom this distinction was primary would naturally tend to group together all species of change, as being on a par.)

Here, then, is one assumption, or set of assumptions, in this part of the argument. I move now to the second of my two observations. Not only does the argument require that “motion” includes “coming to be,” and that anything which is a first principle of “motion” is a first principle of *all* species of “motion”;<sup>18</sup> it also appears, at least at first sight, to require that there is just *one* principle of motion in the entire universe. For the argument for the indestructibility of the principle of motion rests, as we saw, on the assertion that all γένεσις would cease with its destruc-

tion. But if there was more than one principle of motion, it is not clear why the destruction of any *individual* principle would result in the collapse of the universe; so long as at least one principle was in operation at any given time, it seems (given Plato's other assumptions) as if  $\gamma\epsilon\nu\epsilon\sigma\iota\varsigma$  could continue. Apparently, then, Plato is simply taking for granted that there is *not* more than one principle of motion.

But there is something very peculiar about this. Recall that Plato is just about to identify that which moves itself with soul (this is step (1) of the main argument, according to my analysis). Since subargument B has identified that which moves itself as the principle of motion, we are led to infer, by simple transitivity, that soul is the principle of motion. But now, if the logic of subargument B requires that there be just one principle of motion, it appears to follow that there must be just one soul. And this, besides being plainly contrary to any commonsense view of the soul, seems flatly to contradict the myth which immediately follows the proof, in which a plurality of individual souls is discussed. Something has gone wrong here; and since, on this view of the matter, the inconsistency seems so blatant, we should probably conclude that the fault lies not in Plato, but in our interpretation of him.

One way to eliminate this apparent contradiction might be to suggest that Plato is implicitly operating with some notion of a World Soul. That is, perhaps he is thinking of individual souls as being all ultimately aspects of some larger, unitary soul of cosmic proportions. In this way the plurality of individual souls would be only superficial and would be consistent with the principle of motion in the universe being in a deeper sense one, as the argument for imperishability requires. Besides making Plato's train of thought coherent, this proposal seems to have at least some historical plausibility. First, a doctrine of World Soul certainly was adopted by various later philosophical schools—by the Stoics and, perhaps more significantly, by the Neo-Platonists—and it is hard not to see the present passage as in some way prefiguring these later ideas. Second, the *Timaeus* makes large use of the concept of a World Soul (though it looks here as if individual souls exist in their own right, in addition to the World Soul, not as aspects of it<sup>19</sup>). Despite these parallels, however, I do not think this is the right way to understand this part of the *Phaedrus*. It would not be surprising, in view of the parallels, if Plato were operating here with a conception of World Soul; but the parallels do not provide any positive evidence that he is doing so. And the fact is that there is absolutely no hint of any such conception in the *Phaedrus* itself. We find no suggestion, either in the proof of immortality or in the succeeding

myth, that all our souls are ultimately aspects of the same thing, or that our ultimate goal, in striving to escape from the cycle of rebirth, is reabsorption in some larger unity. On the contrary, I would say, it is the individuality of our souls, the differences between them, that is emphasized in the myth.<sup>20</sup>

I conclude that we are not entitled to read the hypothesis of a World Soul into the *Phaedrus*; and the problem remains as to how we are to make sense of Plato's apparent assumption that there is just one self-mover. But perhaps there is another way to understand the matter. Possibly he is using “ψυχή” as a mass term, analogous to “water” or “electricity.”<sup>21</sup> In this case, it would not be that there was one soul—the World Soul; rather there would be a single kind of *stuff* (just as water is a single kind of stuff), of which individual souls (like individual pools of water) consist.<sup>22</sup> Now, anything which is true of water as such (for example, that its chemical composition is  $H_2O$ ), or of electricity as such, is also true of any individual pool of water, or of any individual electrical current; and similarly, if soul is being conceived of along the same lines, anything which is true of soul as such will also be true of individual souls. But the property which, by the end of the argument, we discover to be true of soul as such is precisely immortality. Thus it will follow, on this understanding, that any given individual soul is immortal—a conclusion which Plato obviously wants to uphold. Moreover, if this is right, we can respect both the logic of subargument B and the plurality of individual souls. Sub-rgument B requires, in some sense, the singularity of soul. But this does not mean that we need to ascribe to Plato the view that, out of all the *particulars* in the universe, only one is a soul. If “soul” is a single kind of stuff, the argument will work perfectly well. For if this stuff is responsible for all  $\gamma\epsilon\nu\epsilon\sigma\iota\varsigma$  in the universe, the destruction of this stuff would indeed result in the universe's collapse; and from this it does indeed follow, on Plato's assumptions, that the stuff will necessarily not perish. But all this is quite compatible with saying that there are many individual souls—just as water (or electricity) is a single kind of stuff, yet there are many pools of water (and many electrical currents). Equally, the singularity of the stuff “soul” is quite compatible with the idea that souls have differentiable parts, and with the idea that there are qualitative differences between souls (both of which we subsequently learn in the myth of the charioteers and their horses). For water, electricity, and the like are similarly differentiable; water is composed of hydrogen and oxygen, and different electrical currents come in different voltages.

This understanding of soul seems, then, to fit the context somewhat better than the previous idea of a World Soul. In addition, Plato's usage of the term “ψυχή” itself seems more consonant with the notion of soul as a stuff (in the sense I have explained). Throughout the argument, “ψυχή” occurs in the singular without article. If he had in mind a World Soul, one might have expected him to refer to it using the definite article; but he never speaks of “ἡ ψυχή,” only of “ψυχή.”<sup>23</sup> Grammatically, that is, he does seem to be using the word as a mass term;<sup>24</sup> and this is exactly what we would expect if he is thinking of soul as a stuff. While the evidence is all too nebulous to admit of certainty, I think we are entitled to at least a tentative conclusion—namely, that the underlying picture of soul with which Plato is operating is indeed this latter one, and not the one which involves a World Soul.<sup>25</sup>

I have dwelt at some length on the presuppositions of subargument B, both because of their interest for our broader understanding of Plato (a matter to which I shall return shortly), and because of the great difficulty of elucidating them. The remainder of the argument can be dealt with much more briefly. Our analysis had reached 245e2. The following sentence (e2-4) contains a statement of both the main premises (2) and (1) (in that order). To recall, premise (1) states that *soul* is that which is its own source of motion; and Plato now offers us a consideration in favor of this assertion. The argument is simply that any body with an external source of motion is “soulless” (ἀψυχον), whereas one with an internal source of motion is “ensouled” (ἔμψυχον) (e4-6). On its own, this seems somewhat flimsy; it is not at all clear that it justifies the conclusion that “soul” is a self-mover in the very strong sense used above. We may certainly make a distinction between bodies that need to be pushed or pulled by something else in order to move, and bodies that do not; and these two classes do seem to correspond with the classes of soulless and ensouled bodies respectively. But it is surely quite possible, for all that Plato has said, that bodies of the latter class should have received some *initial* propulsion into motion from outside, even if they need no further propulsion after that. And this, of course, would disqualify them as self-movers; for as we saw, Plato is thinking of a self-mover as a *first principle* of motion.<sup>26</sup> His argument is far from showing that soul is responsible for the motion (including the γένεσις) of *everything else* that moves. His first premise—and with it, a key point in the striking conception of soul the argument for immortality reveals—is more assumed than argued for. This concludes my detailed examination of the argument; for the remaining lines 245e6-246a2 merely contain a restatement of premise (1),

followed by (3), the conclusion of the entire argument.

## II

Let us now shift our approach somewhat. As I said, Plato intends that the argument should convince us by the force of its logic; we are not to take on trust what he says, but to give our assent if and only if we are rationally compelled. This is nothing unusual for Plato—though the density and rigor of this particular passage is unusual; he constantly exhorts us, through the mouth of Socrates, to “follow the argument” where it leads us, and not to base our views on anything else. So far, I have been trying to approach the argument in precisely this spirit, attempting both to understand exactly what Plato is saying, and to determine how much of it we can reasonably accept. For the rest of the paper, I shall adopt a more exclusively historical attitude. As we have seen, the argument contains various assumptions or presuppositions which we might very well question; having engaged with him in the fashion which he himself apparently invites, we find that he fails to convince us of his conclusions. There are no doubt many reasons why this is not surprising, and to discuss them would take us far afield into more general issues in philosophy and in intellectual history. However, I prefer to stick more closely to the argument itself, and to see how it can help to increase our understanding of Plato's thought. The assumptions or presuppositions that I just mentioned are of great interest in themselves. For as I have suggested, they reveal a certain underlying conception of the nature of soul; and it is instructive to compare this conception with other, generally more explicit, conceptions of soul, both in other dialogues and elsewhere in the *Phaedrus*. This is what I now intend to do.

The underlying conception of soul has several features, and we should briefly review them. The most obvious feature, and the one scholars have noticed most often, is that soul, and soul alone, moves itself; indeed, this is simply my premise (1). But Plato's view is stronger than this alone would imply. Recall the feature of soul which we uncovered in examining the phrase ἀτε οὐκ ἀπολεῖππον ἔχει, Plato's support for premise (Ai)—namely, that soul, the self-mover, is in *unceasing* self-motion, and necessarily so. Further, we discovered, in connection with premise (Bi), that soul has a cosmic significance; for as self-mover, soul moves not only itself but everything else—where “moves” is not restricted to what we would call “motion,” but includes (probably among

other things) generation. And finally, there was the puzzling point raised by premise (Bii), that “soul” is being conceived of as in some sense unitary. While the issue is very obscure, we concluded that it was best to understand “soul” as a unitary (nonmaterial) stuff, of which our souls consist. How does a conception of soul containing these features relate to other conceptions that we find in Plato?

I shall begin by summarizing a few familiar points from the *Phaedo* and the *Republic*. In these two dialogues, a strict division is made between two realms, the realm of the changeable and the realm of the changeless. The former is the realm of sensible objects, the latter of intelligible objects; and in the *Republic*, at least, it is the latter class of objects which, just *because* they are changeless, alone fully warrant the title of “things that are” (*οὐτα*). Now, in the *Phaedo* the soul is argued to be akin to the Forms—that is, to the objects granted to belong to the non-composite, changeless, intelligible class; and from this it is deduced that the soul, too, is noncomposite and changeless, and therefore immortal (78b4 ff.). Or at least, it is claimed, the soul *in its essential nature* is noncomposite and changeless. However, because of its association with the body, it is bound, in this life, to partake of the changeable to some extent—though this taint is something we should strive to minimize.

In the *Republic*, the picture of the soul is more complicated; for in Book IV we are told that the soul consists of three parts, and that our task is to achieve the appropriate harmony between them. Now, the two lower parts of the soul, at least, appear to be very much subject to change;<sup>27</sup> and so it might seem that the *Republic*’s view of the soul was a fundamental departure from that of the *Phaedo*. However, a strong case can be made for saying that the two dialogues are really consistent with one another. For one thing, the *ideal* state for the soul in the *Republic* (though most people are incapable of attaining it) is still rational contemplation; and this is a state where the lower two parts of the soul, and change itself, appear to have no place. But more importantly, the argument for immortality in Book X seems to imply that, as in the *Phaedo*, the soul *in its true nature* is noncomposite and changeless. As we see it, it is “composed out of many elements and not enjoying the finest composition” (611b5-6), and is “full of variability and unlikeness and difference” (b2-3). However, this is due to its association with the body; if we were to see the soul as it really is, free from such association, we would probably find that it was just the opposite. It is hard not to read this as suggesting that in its true nature, the soul is not tripartite—that the division argued for in Book IV pertains only to the soul as embodied.<sup>28</sup> To be sure, the tone is

tentative; but the view being expressed seems clear enough.

Let us return to the *Phaedrus*. How does the *Phaedo/Republic* view of the soul compare with the underlying conception we discovered in the *Phaedrus*’ argument for immortality? The first feature in that underlying conception was that soul was that which moves itself; and since “motion” here includes all kinds of change, this appears to stand in simple contradiction with the view I have just been describing. Now, if we restrict ourselves to the bare assertion “soul is that which moves itself” (which was my premise (1)), it is not, in fact, clear that the two views are formally inconsistent. For the *Phaedo* and the *Republic* allow that the soul does undergo change when in the body; and indeed, the evidence for premise (1) concerns the distinction between ensouled and soulless *bodies*. The issue is not, then, as simple as it might seem. However, as we saw, the underlying conception in the *Phaedrus* is not just that souls do at times move themselves, but that soul is necessarily *always* in motion. In other words, soul *must* be in motion both in and out of the body; in the *Republic*’s terminology, the soul “*in its true nature*” is in unceasing motion. And this conception does undoubtedly stand in opposition to that of the *Phaedo* and the *Republic*. Instead, it is aligned with views which we can confidently place at the end of Plato’s career. I am thinking in particular of the *Laws*, where the definition (*λόγος*) of soul is said to be “that motion which is capable of moving itself” (896a1-2); except that soul is here seen not as the *thing which* moves itself, but as *itself* a *species of motion*,<sup>29</sup> this is identical with what the *Phaedrus* refers to as the *οὐσίαν τε καὶ λόγον* of soul (246e3).

But another important passage is *Sophist* 248a4-249d4. Here Plato argues, against the view of the *Republic* and elsewhere, that “that which is” is not restricted to the changeless, but encompasses the changeable as well; in fact, it is plausible to suppose that the putative target of the argument, some people called the “Friends of Forms,” is really his own earlier self. Most interestingly, the key point in the argument is that if the Friends of Forms were right, then “life, soul and thought” (248e6-7) could not belong to the realm of that which is. Obviously it is assumed not only that soul is a “*thing which is*”—which, I take it, Plato assumed at every stage<sup>30</sup>—but also that soul belongs to the class of changing things. This argument, then, is highly congenial to the underlying conception of soul in the *Phaedrus*. For there, as we saw, soul is in constant self-motion; and by the *Phaedo/Republic* criterion for Being, this would entail that soul was not a “*thing which is*.” So far, therefore, the matter appears relatively straightforward, the *Phaedrus* seems to square with

Plato's later thought, not with what is usually seen as his middle period.

Let us now turn from the *Phaedrus*' argument for immortality to the myth which follows. At first sight, this seems to run against the conclusion I just reached. For the myth includes several elements which are clearly reminiscent of the *Phaedo* and the *Republic*. First, the image of the charioteer and two horses, one of which is obedient to the charioteer and the other not, evidently recalls the tripartite soul in the *Republic*; at 247c7-8, indeed, the charioteer is explicitly identified as *νοῦς*. Again, the supreme vision the soul can experience is the "vision" of the Forms; and as in the *Republic*, the Forms constitute the realm of "that which is" (247d3), and are the entities of which true knowledge is possible—as contrasted with the sensible realm, the realm of *γένεσις*, of which one can have only opinion. Moreover, it is through *recollection* of the Forms beheld before birth that understanding in this life is said to be possible (249b-c); this doctrine goes back to the *Phaedo* and to the *Meno*. Finally, of course, as in the *Phaedo*, the goal of life is to escape from the cycle of rebirth as soon as possible, and the means is to live as a philosopher.

We cannot deny, then, that there are many continuities between these other dialogues and the *Phaedrus*.<sup>31</sup> At the same time, though, there are several crucial differences, which make the continuities much less significant than they may seem. One point which has often been noticed is that in the *Phaedrus*, unlike the *Republic* (if my own and others' reading of it is correct), all three parts of the soul are immortal; even the gods' souls have three parts, though in their case the different parts are not in conflict.<sup>32</sup> Again, the end-point of the soul's progress is not changeless and eternal contemplation of the Forms, but an eternal traversing of the heavens, punctuated by contemplation of the Forms at intervals. Moreover, if we take the details of the myth seriously—as I think we should—the lower two parts of the soul (represented by the horses) play an indispensable role in this eternal traversing, and in transporting the charioteer, reason, to the point where it can contemplate the Forms, reason on its own appears to be relatively inert. The ideal state is no longer one where the lower, changeable parts of the soul are transcended. On the contrary, they are just as important as reason itself to the soul's fulfilling of its final destiny; and this final destiny itself consists not of freedom from all change, but of constant, albeit regular, motion. Finally, while the *Phaedo/Republic* conception of a schism between Being and Becoming is maintained, the criterion for the division can no longer be quite the same. For, as I said earlier, I presume that at every stage Plato would have held that the soul is a "thing which is"; but if so, changeless-

ness can no longer be a necessary condition for Being—for as we have seen, the soul, even in its true nature, is no longer changeless. Plato does not explicitly tell us that the soul is an *ōv*; indeed, he seems to restrict the title to the Forms. However, he does say that the soul "feasts on" and "is nourished by" *τὸ ὄν*; and the metaphor of nourishment must, I take it, imply that the nature of the soul is not basically alien from that which nourishes it—in other words, that the soul, too, is an *ōv*.

On this last point, it may be that Plato did not manage to take account of all the implications of what he was saying. But even so, there is at least a shift away from the *Phaedo* and the *Republic* and toward the *Sophist*, where it is argued that we must not restrict the title of *τὸ ὄν* to changeless entities. As for the other issues I just mentioned, the *Phaedrus* myth seems to be definitely aligned with the view of soul propounded in the *Laws*, and against those earlier dialogues. As we saw, the *Laws* defines soul as self-motion; and the souls in the *Phaedrus* myth, instead of being ideally static, are in constant, self-propelled motion. Moreover, the myth assigns an essential role in this motion to those parts of the soul other than reason. Again, in the *Laws*, the following list of motions of the soul is given: "wish, reflection, foresight, counsel, judgement, true or false, pleasure, pain, hope, fear, hate, love" (897a1-3, A. E. Taylor's translation). The *Laws* does not make explicit use of the notion of a tripartite soul. However, if one approaches this list with the tripartite soul in mind, and asks which items on the list can be assigned to the rational part, the answer seems clearly to be "not all of them": some, at least, would have to belong to the other two parts.

So it turns out, I think, that the *Phaedrus* myth is in central respects closer to Plato's later views than to the *Phaedo* and the *Republic*; the similarities with the latter dialogues seem to me less significant than the differences. We can now see, too, that the myth is, after all, in agreement with the argument for immortality which immediately precedes it. Given the similarities between the myth and the earlier dialogues, it looked as if the myth and the proof were at odds with one another. In fact, however, both view the soul as being in eternal motion; and this sets them both apart from the earlier dialogues, where the soul is in its true nature changeless, and alongside certain later dialogues.<sup>33</sup> There is, then, a deep connection between the proof and the myth—a connection which has rarely been noticed.

I have so far said nothing about the place of the *Timaeus*. There is, of course, much in the *Timaeus* about the soul; and since it is usually considered a late dialogue, one might expect that it, too, would be naturally

grouped with the *Phaedrus*, *Sophist*, and *Laws*. Now, the dating of the *Timaeus* has been the subject of much heated debate, and I cannot attempt to deal with this thorny question here. However, it is striking that, with regard to the matters we have looked at so far, the *Timaeus* is in several respects closer to the *Phaedo* and the *Republic* than to the later group. First, the *Timaeus*, like those two dialogues, makes a strict division between the realms of Being and Becoming, with changelessness (including motionlessness) repeatedly cited as a feature of the former realm and change as a feature of the latter.<sup>34</sup> This clearly separates the *Timaeus* from the *Sophist* and the *Laws*, and separates it to some extent from the *Phaedrus*—though, as we saw, the *Phaedrus* is not entirely straightforward about the relation between change and Being. Second, the *Timaeus* employs the notion of a tripartite soul; however, here Plato is quite explicit that only the rational part is immortal, which places it with the *Republic* and against the *Phaedrus*. Again, at 69c8 ff. there occurs a list of “terrible and necessary affections” associated with the mortal parts of the soul; these are pleasure, pain, rashness, fear, anger, hope, sensation, and love. The immortal part has none of these, except insofar as it is polluted by the mortal parts; in fact, it is to minimize this pollution that the mortal parts are housed in separate parts of the body.<sup>35</sup> But this list has much in common with the list of “motions” of the soul in the *Laws*, which I quoted a little while ago; and those “motions” are certainly not regarded as merely temporary, or as a pollution of any kind. Once more, the *Timaeus* appears to fit less with the later dialogues than with the earlier ones. Finally, the *Timaeus* does not define soul as that which moves itself, or suggest that it is responsible for the motion of everything else. It is true that the World Soul does undergo certain harmonious cosmic revolutions (35a1 ff.); but these are initiated by its creator, not by itself, and they are not what causes motion in other things—though they do appear to play some kind of regulating function.<sup>36</sup>

What we seem to have found so far is that there are roughly two sets of ideas about the soul in Plato—one of which appears to be earlier and one later—and that the *Phaedrus*, both the proof and the myth, displays essentially the later set. Now, it may look as if the introduction of the *Timaeus* does not importantly change this picture. For my remarks in the last paragraph might suggest that we could retain the twofold division, and simply group the *Timaeus* with the earlier set. There are some oddities, to be sure; but the *Timaeus*’s view of the soul, it might be said, is basically the same as that of the *Republic* and the *Phaedo*. However, the

situation is not so simple. Recall that the underlying conception of soul that I detected in the *Phaedrus*’ argument for immortality had several features. Two of these features we have not yet discussed; one was that soul plays a vital role in the operation in the cosmos, and the other was that soul is in some sense one—probably a single kind of stuff. These features complicate the picture in two ways. First, they alert us to the fact that what I have been calling the later conception is actually a lot less unified than I have implied. And second, they incline us to group the *Timaeus* with the later set of dialogues, and not, as we were doing just now, with the earlier set. I shall briefly expand upon these points.

In the *Phaedrus*’ proof, Plato regards soul as having a cosmic significance; for it turns out that soul is responsible for the motion, including the  $\gamma\acute{e}ne\sigma\iota\varsigma$ , of everything that is not soul. This agrees with the myth; at 246b6 we are told that “all soul takes care of all that is soulless.” Now, this idea is not present in the *Phaedo* or the *Republic*. In both dialogues there is the suggestion that the universe is ordered for the best, and this may suggest some beneficent ordering force. But Plato does not tell us about any such ordering force, and he certainly does not suggest that it is soul. Again, the *Phaedrus* marks a departure from those dialogues; and in this respect, again, it is in agreement with the *Laws*. As the “motion which moves itself by itself,” soul is said in the *Laws* to “direct everything in heaven, earth and sea” (896e8-9). But the other dialogue which makes soul a cosmic force is the *Timaeus*. As we saw recently, the view is not the same as that of the *Phaedrus* and the *Laws*. It is not suggested that the World Soul initiates motion in other things, nor that it is self-moving. In the framework of the *Timaeus*, the origin of motion can be accounted for in other ways; there is the Craftsman, who creates and sets in motion the World Soul itself, and there is also the disorderly precosmic motion in the Receptacle of Becoming. Instead, the World Soul is said to be the “mistress and ruler” (34c5) of its body, the cosmos; and its function is apparently in some way to preserve order. A further difference (if I was right earlier) is that the *Phaedrus* does not employ the notion of a World Soul. (The position of the *Laws* in this respect is not clear to me, for reasons which I shall go into shortly.) But despite these differences, the *Timaeus* seems in at least one way closer to the later dialogues than the earlier ones.

The final feature of the conception of soul we extracted from the *Phaedrus*’ proof—albeit tentatively—was that soul is a single kind of stuff (in the sense explained—see note 22); how does this compare with other dialogues? The *Phaedo* and the *Republic* contain no hint of such a

conception, though they do not obviously exclude it. The one other dialogue where soul definitely is regarded as a stuff in my sense, and a single one, is the *Timaeus*. As we have noted, the Craftsman mixes a certain compound substance, and out of this substance all souls are formed. Now, the substance put together in the mixing bowl clearly qualifies as a stuff; this stuff could very well be designated by the mass term "soul" (though Plato does not appear to do so). Individual souls (both the World Soul and human souls) consist, then, of portions of the stuff "soul"; and this is precisely the picture I found to be suggested by the *Phaedrus*. One might say that this is a far too literal reading of the *Timaeus'* mythical apparatus. But I believe that one should always take seriously the details of Platonic myths. Obviously some features, such as the mixing bowl, will be factored out when we try to penetrate to the core of doctrine Plato is trying to present. But if Plato had not meant to convey the impression that there is a kind of stuff of which souls are composed, he could have written a different myth that did not have these implications; he was not forced to write it as he did. (On this point, see also notes 3 and 32.)

Here, then, is another apparent point of contact between the *Phaedrus* and the *Timaeus*.<sup>37</sup> Let us see, finally, whether we can relate the *Laws* to these two. Plainly, the evidence for any notion of soul as a stuff in the *Laws* will be indirect; but there is one point that seems to me suggestive. This is that Plato is strikingly evasive about the *number* of souls that run the cosmos. He is clear that this is done by one or more *good* souls; but he refuses to commit himself as to how many. (This is why it is hard to tell whether or not he has in mind a World Soul.) Mostly he uses the singular *ψυχή* without article, as in the *Phaedrus'* proof; but he speaks also of "soul or souls" (899b5, cf. 898c7-8), and he never decides whether the singular or the plural is more appropriate. He seems, then, to be either uncertain about, or uninterested in, the individuation of souls; yet this appears not to detract in any way from his confidence in the doctrine of soul as activator of the cosmos. Now, this does not, of course, show conclusively that he is thinking of soul as a stuff. On the other hand, if he was conceiving of soul (or simply of good soul) as a uniform stuff distributed about the universe, this insouciance about the number of good souls would be entirely understandable. If there is a certain kind of stuff whose nature it is to activate the cosmos, then it is indeed of minor importance—supposing that the motions of the cosmos are what we are seeking to understand—how portions of this stuff are to be individuated. However, this is as far as we can pursue the question. It is possible that the *Laws* as well as the *Timaeus* is in line with the *Phaedrus* on this

matter; but the evidence is too slender for us to decide with any confidence. Moreover, even if all this is correct—and I am far from insisting on it—it still leaves out of account the *Laws'* idiosyncratic notion of *evil* soul. The one thing Plato is clear about in the *Laws* with regard to the number of souls, is that there are at least two. The soul or souls that activate the cosmos are good; however, there must also be one or more evil souls (896e4-6), to account for the presence of evil in the world. Of course, it might still be that there was a single stuff "soul" (whose essence was self-motion) encompassing both good and evil soul. But it is obviously impossible to tell whether or not this is Plato's conception; the remarks on evil soul—indeed, on soul in general—in the *Laws* are just too sketchy.

What general conclusions, if any, can we draw from this investigation? It is sometimes suggested that the *Phaedrus* is a "transitional" dialogue. Now, we have examined only a tiny fraction of the dialogue; and it may well be that, concerning topics which we did not touch upon, this verdict is a fair one. But with regard to the nature of the soul, I believe, it needs at least to be substantially qualified. For to say that the dialogue is transitional is to imply that there is an earlier set of doctrines and a later set of doctrines, and that the *Phaedrus* is between the two. This is not entirely devoid of truth, but it is misleading in two ways. First, as we have seen, in as much as one can speak of two groups of dialogues, an earlier and a later, the *Phaedrus* seems in most respects (again, of those that have been the subject here) to belong solidly with the later group; the transition, for the most part, occurs before the *Phaedrus*, not during it. But second, there are not just two views of the soul in Plato. The *Phaedo* and the *Republic* may with some justice be considered as expressing a single view—though even this is perhaps an exaggeration; but among the later dialogues, there is far less homogeneity. In particular, the position of the *Timaeus*, on this as on so many other issues, is baffling. It is dangerous, then, to speak of "transitions" in this context. We can certainly find various trends in Plato's writings on the soul; and we can try to describe the position of the *Phaedrus* among these trends. But we should not think in terms of stable bodies of doctrine between which the *Phaedrus* might be considered as intermediate. This conclusion is no doubt nothing surprising; people have often told us not to force Plato into a straitjacket. Much more interesting—or so I hope—are the detailed considerations which have led us there.

## Notes

1. *Hermeiae Alexandrini in Platonis Phaedrum Scholia*, ed. P. Couvreur (Paris: É. Bouillon, 1901) ad loc.

2. The most detailed discussions I am aware of are in R. Hackforth, *Plato's Phaedrus* (Cambridge: Cambridge University Press, 1952), 64–68, and T. M. Robinson, “The Argument for Immortality in Plato's *Phaedrus*,” *Essays in Ancient Greek Philosophy*, ed. J. P. Anton and G. L. Kustas (Albany: State University of New York Press, 1971). But both of these, while they do cover the main steps of the whole argument, deal with most of them extremely briefly; in addition, Robinson's analysis seems to me mistaken in several places. Among other writers and commentators on the *Phaedrus*, W. J. Verdenius (“Notes on Plato's *Phaedrus*,” *Mnemosyne Series* 4, 8 [1955]: 265–89) and G. J. de Vries (*A Commentary on Plato's Phaedrus* [Amsterdam: Hakkert, 1969]) offer only isolated remarks, mostly on passages where the text is in doubt. I. M. Crombie (*An Examination of Plato's Doctrines* [London: Routledge & Kegan-Paul, 1962]), in a long and exhaustive chapter on Plato's views on the soul (Vol. 1, ch. 7, “The Philosophy of Mind”), devotes only a page and a half (325–27) to the argument, mostly at a very abstract level; and Martha Nussbaum, in a discussion of the *Phaedrus* which covers almost every major topic in the dialogue (“‘This Story Isn't True’: Poetry, Goodness and Understanding in Plato's *Phaedrus*,” in *Plato on Beauty, Wisdom and the Arts*, ed. Julius Moravcsik and Philip Temko [Tottowa, N.J.: Rowman & Littlefield, 1982]), dismisses it in literally a sentence (106–107).

3. This contrast should not be taken to imply that the proof is more important than the myth, or that the myth is not “real philosophy.” In this connection, it is interesting that Plato uses the word ἀπόδειξις (245c1, 4; cf. ἀποδεικτέον, b7) to refer not simply to the proof of immortality, but to the entire ensuing discussion; the “demonstration” is of the fact that the madness of the lover is divinely inspired, and so encompasses the myth as well as the proof. In Plato's view, then, myth and proof are equally valid ways of showing things—though of course, each may be appropriate in different circumstances.

4. Some (e.g., de Vries, *Plato's Phaedrus*, following Denniston) have likened the style of the proof (in addition to its thought—but that is another matter) to that of certain Presocratics. The similarity may be there, and may be intentional; but this does not, I take it, detract from the plausibility of my suggestion here—there may be more than one reason for his adopting the style that he does. Raphael Demos (in “Plato's Doctrine of the Soul as a Self-Moving Motion,” *Journal of the History of Philosophy* 6 [1968]: 133–45) goes so far as to suggest, largely on the grounds of this stylistic difference, that the argument is a later insertion. This seems to me quite unnecessary; it is not at all unusual for Plato to change his style radically between adjacent passages.

5. This analysis of the argument's broad outline agrees most closely with

that of Hackforth (*Plato's Phaedrus*). The only point at which I would take issue with him is that he regards the second of the two arguments for (2) as subordinate to the first; it seems to me that the two are parallel and equally important. Hermias' reconstruction is also roughly in agreement with my own (see 104.4–12 for his introductory statement of the premises). He thinks that the argument as a whole divides into two; in his view, the first subargument runs (using my symbolism) (1), (Ai), (Aii), therefore (3), and the second (roughly) (1), (Bi), (Bii), therefore (3). (This is not quite right, since he further subdivides my (Bii), and his view of the logic at this point is slightly different from mine; on this, see further note 14.) The main difference, then, is that he does not include Plato's statement of my main premise (2); but this difference is not important, except from a strictly formal point of view. The analysis of Robinson (“The Argument for Immortality”) is rather more distant from mine. Partly for this reason, it would be a somewhat arduous task to criticize it directly; I prefer simply to offer my own rival interpretation, and hope that it prevails on its own merits.

6. There has been much division of opinion over the alternative readings ἀεικίνητον and αὐτοκίνητον in 245c5. It seems to me that the matter cannot be decided on the basis of their relative appropriateness to the argument as a whole; either reading would yield a closely knit logical progression. It is true that the reading αὐτοκίνητον makes for some repetition; but this would be quite tolerable—it by no means suffices (as J. B. Skemp thinks—see *The Theory of Motion in Plato's Later Dialogues* [Cambridge: Cambridge University Press, 1942], 3, n. 2) to rule the reading out. Hackforth defends ἀεικίνητον on the grounds that it gives us a statement of an ἔνδοξον—that is, a premise which could be expected to find general acceptance; but against this, Ackrill (*Mind* 62 [1953], 278) seems to point out that an argument does not need to begin with a statement of some premise that is uncontroversial. Logical considerations aside, however, the evidence for ἀεικίνητον seems to me entirely superior. The mss. agree on ἀεικίνητον; αὐτοκίνητον is found only in one Oxyrhynchus papyrus (1017). In addition, Hermias clearly had ἀεικίνητον in front of him, and so did Cicero when he translated the passage for the *Somnium Scipionis* (a section which he also quotes in *Tusculan Disputations*, I. 54); the Latin is “quod semper movetur, aeternum est.” For a full discussion of the issue, which supports this general position, see F. Decleva Caizzi, “AEIKINTON o AYTOKINTON?”, *Acme* 23 (1970): 91–97.

7. Thus c5–7 (τὸ δ' ἄλλο κινοῦν, etc.) seems to me to look forward to the following sentence, rather than back to τὸ γὰρ ἀεικίνητον ἀθάνατον, despite the evident parallelism between the pairs ἀεικίνητον/ἀθάνατον and παῦλαν κινήσεως/παῦλαν ζωῆς. Burnet's punctuation suggests the opposite. (Translations, here and elsewhere, are my own.)

8. If this were the line of thought that was motivating Plato—and it is my best conjecture on the subject—then it would in one respect anticipate Aristotle's ideas on self-motion. For on this conjecture, Plato is led very naturally

into thinking of self-movers as consisting of (at least) two parts, the active and the passive part, so to speak. Aristotle's analysis of the concept of self-motion (*Physics* Bk. VIII, ch. 5) makes central use of just such a division; a thing which moves itself must, he thinks, consist of a part which is moved and a part which causes this motion. Aristotle argues convincingly that this latter part cannot itself be in motion, which leads him to the concept of the *unmoved mover*; it turns out, then, that, contrary to Plato, the ultimate cause of motion is not a self-mover. However, a picture in which Aristotle developed beyond Plato, having begun by holding ideas similar to his, is presumably one we are bound to adopt in any case; and the development would perhaps be a little smoother if Plato's views were of the type I suggest. But of course, all this is pure speculation.

9. That there is an error common to the two passages is noted in passing by Demos ("Plato's Doctrine," 135). The standard view of the relation between them seems to be that they are arguments of essentially the same kind, but that the argument in the *Phaedrus* is of a more empirical nature; see Hackforth, *Plato's Phaedrus*, 68 and Robinson, "The Argument for Immortality," 347. I agree about the similarity, but I cannot see that the *Phaedrus* argument is more empirical. It is true that the *Phaedo* argument makes use of the terminology of Forms; but as far as I can see, this is not really essential to it. The crucial point is simply that the soul necessarily brings life. In the *Phaedrus*, the crucial point is that the soul is necessarily in constant self-motion. Each argument rests, then, on a very basic feature of the soul, and I do not see that one of these features is any more empirical, or observable, than the other. It is also true that the *Phaedrus* argument owes more to Presocratic thought (on which see also Skemp, *The Theory of Motion*, 3-10); but to equate "Presocratic" with "empirical" (as Robinson, at least, appears to do) seems to me highly inadvisable.

10. For a similar use of *αεί* and *οὐποτεί*, consider one of the ways in which Plato commonly contrasts Forms and particulars. Forms are frequently said to be "always" the same; each Form "always" possesses the property which makes it the specific Form that it is. Particulars are sometimes one way, sometimes the opposite way: but the Form of Beauty, for example, is "never" other than beautiful. Examples of this usage are *Rep.* 479a2-3, *Pr.* 74b7-c3, and *Symp.* 211a3, b1-2.

11. Curiously enough, Plato appears to recognize the error in the *Phaedo* immediately after he has committed it. (Curiously, because the interlocutors are made to regard the argument we have been discussing as entirely conclusive—105e9.) He continues by pointing out that it needs to be shown that the soul is imperishable. He appears to be saying that to establish that the soul "does not admit death"—i.e., that nothing is both dead and a soul—is not to establish that it never ceases to exist; and this is essentially the point that I have just been making. (See in particular the paragraph 106b1-c7.) However, his way of closing this gap is rather lame; he simply says that *of course* that which is *ἀθάνατος* is also imperishable (106d2 ff.). If "the soul is *ἀθάνατος*" means only that the

soul does not admit death, in a sense which is compatible with the soul's ceasing to exist—as he appears to have just conceded—then this is plainly begging the question.

12. Indeed, this sentence is dense to the point of being very confusing; it contains a premise for subargument A—as well as a reason for it—and a premise for subargument B; moreover, the same words *μόνον δὴ αὐτὸς κινοῦν* serve as subject in the statement of both. I was originally tempted to think that what I am now calling (Bi) was a second supporting consideration, parallel with *ἄτε οὐκ ἀπολεῖπον ἔχων*, in favor of (Ai). But while this would reduce the number of different things going on in the sentence, it cannot be the right way to read it; for no mention is made, here or anywhere, of the first principle of motion being *always in motion*—as would be needed if the assertion in question (that that which moves itself is a first principle of motion) were to constitute support for (Ai). Instead, this assertion leads directly into claims about an *ἀρχή*'s freedom from *γένεσις* and *φθορά*—claims leading to what I now call (Bii); the idea of constant motion, which is at the center of subargument A, is left behind.

13. Reading *ἔξ ἀρχῆς* in d3, and taking *πᾶν τὸ γιγνόμενον* as the subject of *γίγνοιτο*; in this I follow Verdenius and de Vries ("Notes on Plato's *Phaedrus*," and *A Commentary on Plato's Phaedrus*). Buttmann's conjecture *ἔτι ἀρχή* would make for much better sense (and for a closer fit with Cicero's translation), if we could understand *γίγνοιτο* as equivalent to *εἴη*; but this is surely impossible.

14. The balance of opinion now seems in favor of the mss. *γένεσιν* in e1 (also found in Hermeias—see 117.24), and against Burnet's reading *γῆν εἰς ἐν* (following Philoponus). It is generally agreed that *γένεσιν* can be taken as equivalent to *τὰ γιγνόμενα*, and this gives good sense.

Hermeias' view of the logic at this point is puzzling. He says that once it has been shown that a first principle is ungenerated, that by itself suffices to show its imperishability (117.17). In his view, therefore, the considerations offered in d4-e2 form a supplementary, and strictly speaking superfluous, argument *εἰς ἀδύνατον* (117.17-25; cf. 103.11.3-6, where he says that the argument as a whole consists of two syllogisms plus one further argument *εἰς ἀδύνατον*). I fail to understand this; it seems to me that d4-e2 is integral to the main argument, and builds very directly upon the demonstration just above that an *ἀρχή* is *ἀγένητος*. Robinson ("The Argument for Immortality") claims agreement with Hermeias on this issue, but seems to me also to misconstrue Hermeias' view as to the extent of the argument *εἰς ἀδύνατον*, locating it from d8-e2 only.

15. The best comment on this point may be Hackforth's (*Plato's Phaedrus*, 66-67), who simply says "That was a possibility never contemplated by any Greek thinker." As he points out, the possibility in question is that there should be no *γένεσις* whatever, not that some *particular* cosmos should perish. The

Greeks had no trouble conceiving this latter possibility; Plato himself suggests in the *Timaeus* (41a) that it is possible (though in fact, since the creator is good, it will not happen to this cosmos), and many Presocratic and Hellenistic schools held that it actually occurred.

16. The *Theaetetus* is also relevant here; in the context of the “secret doctrine” associated with Protagoras, Heraclitus, and others, the words *πάντα γίγνεται* and *πάντα κινεῖται* appear to be intended as equivalent (compare 152d8 with 183a5).

17. Hermeias seems quite unruffled by this idea; in discussing this part of the argument, he simply remarks in passing *ἢ γὰρ γένεσις κίνησις ἔστιν* (116.13), as if this is entirely commonplace.

18. This is perhaps too strong. Strictly, all that the argument requires is that anything which is a first principle of *any* species of “motion” is also a first principle of *γένεσις*.

19. The creation of individual souls (or rather, of those souls and parts of souls that are immortal) is described at 41d ff. The creator is said to fashion these souls out of the remainder of the elements previously used to fashion the World Soul, mixing them in the same bowl, and in the same manner (except with some dilution), as was used for the World Soul. Obviously, it is hard to know how, if at all, this is to be cashed out literally. But it is at least implied, I take it, that individual souls are not simply aspects of the World Soul; they enjoy a separate existence.

It is sometimes said that a World Soul is also to be found in the *Laws*, but this is much less clear; as far as I know, a World Soul is never explicitly referred to in that dialogue. I shall briefly touch upon this point again near the end.

20. Think, for example, of the distinction between gods’ souls and human souls (246a-b); between the twelve companies of souls, each led by a different god (246d6-247a4); and between the ten types of human lives, corresponding to various conditions of human souls at birth (248c2-e3). While none of these distinctions is formally inconsistent with the notion of a World Soul, they do indicate that Plato is choosing to focus on the differences between souls, not on their fundamental unity; and this, I think, makes it implausible to suppose that he has a World Soul in mind—given that there is nothing whatever in the text to suggest it.

21. In what follows, I am expanding on a suggestion by Tony Long. This is perhaps an appropriate place also to thank him more generally for much helpful criticism and encouragement as this paper was developing—especially on the matters I am now about to introduce.

22. My use of the term “stuff” is perhaps a little unusual, and I should define it as precisely as I can. To begin with what is uncontroversial, all stuffs are designated by mass terms, but not all mass terms are the names of stuffs. For abstract nouns, such as “hunger” or “wisdom,” qualify as mass nouns (they do not occur in the plural, nor can they be qualified by numerical adjectives); yet it

would be bizarre to call hunger or wisdom a stuff. Now, this might suggest that stuffs are by definition *material*; and in this case I would not, of course, want to suggest that Plato thinks of souls as a stuff. But however it may be in ordinary usage, I intend to use the word “stuff” in such a way that there can be said to be nonmaterial stuffs. In my usage, the referent of any mass noun which is not an abstract noun will count as a stuff. (Here I am assuming some intuitive sense of what is an abstract noun; I admit that it is very hard to give a foolproof definition. On this and several other complications in the notion of a stuff, see V. C. Chappell, “Stuff and Things,” *Proceedings of the Aristotelian Society* 71 [1970-1] 61-76—an article from which I have profited in writing this.) Thus electricity, though not in any obvious sense material, is a stuff; and so, too, is soul—if it is true that Plato is using “ψυχή” as a mass term. For this reason, electricity is perhaps the most useful analogy to keep in mind in what follows.

Another parallel which may be helpful is the concept of “spiritual substance” in early modern philosophy; a famous passage which makes use of this concept is the chapter on identity in Locke’s *Essay* (Bk. II, ch. XXVII, “Of Identity and Diversity”). Here Locke raises the possibility that some contemporary person’s soul might be composed of the same spiritual substance as that which made up the soul of Socrates (though this possibility is, according to him, irrelevant to questions about identity)—just as two material objects existing in nonoverlapping periods of time might be composed of the same material substance. In my usage, we might say that spiritual substance, like Plato’s soul, is a kind of *immaterial stuff* of which souls are composed (though I do not want to take the parallel any further than that); while we may find such a picture highly alien, Locke’s account (even if it is not offered in an entirely serious spirit) does, I think, illustrate that the concept of an immaterial stuff is a perfectly coherent one.

23. The one possible exception is 246b6, at the beginning of the myth, where some mss. have *πᾶσα ἢ ψυχή*, also *ἢ ψυχὴ πᾶσα*. As far as I can gather from LSJ and from Greek grammars, both these phrases would have to mean “The whole soul.” And since it is clear from the context that the reference is not to individual souls (for no individual soul, whether in its entirety or not, “takes care of *all* that is soulless”), it would have to be to a World Soul. If so, however, it would be the sole explicit mention of the World Soul in the whole dialogue; and its intrusion here, after a complicated proof of the immortality of soul which makes no explicit use of such a notion, would, I think, be very surprising. My claim regarding the usage of *ψυχή* should not, therefore, be upset by these ms. variants; we should follow Hermeias and Simplicius, and read *ψυχὴ πᾶσα*. (All modern editors of whom I am aware do indeed adopt this reading.) Given the prevalence of doctrines of a World Soul in later antiquity (including, as I said, Neoplatonic doctrines), one might in any case argue that *ψυχὴ πᾶσα*, which does not entail any notion of a World Soul, is the *lectio difficilior*.

24. This is not necessarily true of *ψυχὴ πᾶσα* in 245c5; taken in isolation,

the phrase could equally be translated as "all soul" or "every soul," and commentators have sometimes exercised themselves over which sense Plato intends. Plainly this cannot be decided by inspection of the phrase itself. The question must be whether one or the other sense is required by the logic of the argument; and this question we are now in a position to tackle. Given the presuppositions of subargument B that I have unearthed, the collective sense "all soul" seems clearly more relevant. As we saw, the principle of motion there under discussion is thought of as in some sense single and whole (whether or not I am right about the exact conception); and the distributive sense "every soul" would run counter to this. On the other hand, the argument is presumably intended to show that each individual soul is immortal; so that the sense "every soul," while less suited to the tenor of the argument itself (and to the usage of "ψυχή" elsewhere in the argument), cannot be decisively ruled out. Hackforth may be right to conclude (*Plato's Phaedrus*, 64) that "the distinction between collective and distributive senses is not here before his mind"; see his discussion for several other useful insights.

25. I should perhaps briefly allude at this point to Hermeias' views on how Plato is conceiving of soul—since it was with his advice on this subject that I began. While he cites Posidonius as having thought that Plato had in mind the World Soul (apparently *as distinct from* individual souls), he himself rejects the idea. Relying heavily on θείας τε καὶ ἀνθρωπίνης in 245c3, he says that Plato must be thinking of the rational soul—since this is the kind of soul possessed by gods and humans, but not by other ensouled beings (102.15 ff.). This seems to me acceptable as far as it goes; but as the last few pages should have made clear, I do not think it goes far enough. Plato's logic raises deeper problems, which I have tried to address; neither Hermeias nor anyone else seems to have faced these problems. Robinson ("The Argument for Immortality") follows Hermeias, but then makes matters worse by suggesting that Plato is perhaps thinking not of the rational soul in *toto*, but of the rational *part* of the soul. This seems to me obviously wrong in view of the myth, where the immortal soul is clearly portrayed as tripartite; more on this in section II.

26. Here it is instructive to compare Aristotle's apparent vacillation on the issue of what it is for something to be a self-mover. On this, see D. J. Furley, "Self-Movers," *Aristotle on Mind and the Senses*, ed. G. E. R. Lloyd and G. E. L. Owen (Cambridge: Cambridge University Press, 1978).

27. This is not, as far as I know, explicitly stated. But there is much emphasis, in the description of the inferior states and inferior souls in Bks. VIII and IX, on the progressively greater instability and variability of the various kinds of life. The less the influence exerted by the rational part of the soul, and the greater the power of the two lower parts (particularly ἐπιθυμία, which is throughout seen as standing in starker opposition with the rational), the more the person, or the state, is subject to change. Also relevant is the discussion, again in Bk. IX (580d-588a), of the different types of pleasures associated with each part

of the soul. Here the pleasures of ἐπιθυμία are identified as the pleasures of the body, and as such intrinsically changeable; by contrast, the pleasures of the rational part come from its communion with that which truly is—that which is changeless—and so are themselves not subject to change. Plato has some trouble fitting the θυμοειδές into this discussion, but he does place it in the same category as the ἐπιθυμητικόν (586c7-d2).

28. The disclaimer at the beginning of the Bk. IV argument, to the effect that the present methods are inadequate for a precise understanding (435c9-d5), may be intended to accommodate this point.

29. This difference is noted by Demos, "Plato's Doctrine," 136.

30. Except in the *Timaeus*. For the *Timaeus'* very peculiar view of the composition of soul, see further note 36.

31. My wording here and elsewhere assumes the standard view that the *Phaedrus* is later than the *Phaedo* and the *Republic*; this seems to me justified by precisely the kinds of comparisons with other dialogues to which this entire half of the paper is devoted.

32. Pace Guthrie, who thinks that the soul is tripartite only during the cycle of rebirth. (See "Plato's Views on the Nature of the Soul," reprinted in Vol. II of *Plato*, ed. Gregory Vlastos [New York: Anchor Books, 1971]; and Vol. IV, 421-25 of Guthrie's *A History of Greek Philosophy* [Cambridge: Cambridge University Press, 1975].) Because the gods' souls are free from internal conflict, he thinks that they must be unitary; and the same goes, he believes, for those originally human souls who have escaped the cycle. In this way he is able to hold that the *Phaedrus* is fundamentally consistent with the *Republic*. That the gods' souls consist of a charioteer and two horses he puts down as just part of the machinery of the myth. This seems to me quite untenable. It is true that Plato switches to the mythical mode here because, as he says, it is beyond human powers to give a fully accurate literal description of the soul; the picture conveyed, therefore, is only tentative and only partial. But this does not mean that we are free to discard any element of the myth that we choose. He composed the myth as he did because he felt that this was the closest approximation to the truth of which he was capable. If he had meant to suggest that the gods' souls did not consist of three parts, he could very easily have composed that part of the myth differently. Nor, incidentally, does he ever suggest that the black horse of the human soul becomes white on the soul's escape from the cycle; rather, it becomes progressively more obedient to the charioteer. Guthrie's interpretation thus involves a further departure from what Plato actually tells us.

Another author who seeks to show that Plato's views on the composition of the soul remain essentially the same is Robert W. Hall, "Ψυχή as Differentiated Unity in the Philosophy of Plato," *Phronesis* 8 (1963): 63-82. As the title suggests, Hall argues that in the *Republic*, the *Phaedrus* and every other dialogue that deals with the subject, the soul is viewed as a differentiated unity. This seems to me almost trivially true, but obviously inadequate to show that Plato's

views did not change. For the question now simply becomes “*into what aspects* is this unity differentiated?” And to this question, as Hall seems quite ready to admit, different dialogues return different (and incompatible) answers.

33. For a fuller discussion of ways in which the *Phaedrus* myth departs from the *Phaedo* and the *Republic* with regard to the nature of the soul, see Nussbaum, “This Story Isn’t True” section II (“Moral Psychology”), 92-107.

34. See, for example, 27d5 ff., 29b5 ff., 35a1-3, 48e5 ff., 51e6 ff.

35. It is curious, however, that at 42a-b most of the items on this list are apparently assigned to the immortal part of the soul also. But Plato does say that this is what happens to it when it is “implanted of necessity in bodies” (a3-4), and that the goal should be to overcome these states as much as possible. Perhaps, then, this is a reference to the pollution by the mortal parts that we find in the other passage; if so, however, it is a confusing one, since the mortal parts have not yet been introduced.

36. Given the *Timaeus*’ criteria for Being, the fact that the World Soul is in motion should mean that it is not in the full sense a “thing which is.” This is indeed the case; for the mixture out of which the World Soul is formed (as well as other souls) is a mixture of τὸ ὄν and τὸ γνώμενον (35a1-3). As far as I know, this idea is perfectly unique in Plato.

37. I have throughout been cautious about attributing the notion of soul as a stuff to the *Phaedrus*; and it will be remembered that the alternative, which I considered but found less likely, was that the *Phaedrus* presupposes a notion of World Soul. It is worth pointing out that even if I was wrong in my choice between the alternatives—even if we should read the *Phaedrus* as implying a World Soul—the closest parallel would still be the *Timaeus*.

## Selected Bibliography

Ackrill, J. L. Review of Hackforth, *Plato’s Phaedrus*. *Mind* 62 (1953): 277-79.

Adam, J. *The Republic of Plato*. Cambridge: Cambridge University Press, 1979.

Annas, Julia. *An Introduction to Plato’s Republic*. Oxford: Clarendon, 1981.

—. *Platonic Ethics, Old and New*. Ithaca, N.Y.: Cornell University Press, 1999.

Apolloni, David. “Plato’s Affinity Argument for the Immortality of the Soul.” *The Journal of the History of Philosophy* 34 (1996): 5-32.

Archer-Hind, R. 1881. “On Some Difficulties in the Platonic Psychology.” *The Journal of Philology* 10 (1881): 120-31.

—. *The Phaedo of Plato*. New York: Arno Press, 1973.

Barnes, Jonathan. Critical notice of D. Gallop, *Plato: Phaedo* (Oxford: Clarendon, 1975). *Canadian Journal of Philosophy* 8 (1978): 397-419.

Bluck, R. S. “The *Phaedrus* and Reincarnation.” *American Journal of Philology* 79 (1958): 156-64.

—. *Plato’s Phaedo*. London: Routledge & Kegan-Paul, 1955.

Bostock, David. *Plato’s Phaedo*. Oxford: Clarendon Press, 1986.

Bobonich, Christopher. “Persuasion, Compulsion, and Freedom in Plato’s *Laws*.” *Classical Quarterly* 41 (1991): 365-87. Rpt. in *Plato II: Ethics, Politics, Religion and the Soul*. Ed. Gail Fine. Oxford: Oxford University Press, 1999, 373-403.

Bolton, Robert. “Plato’s Distinction between Being and Becoming.” *Review of Metaphysics* 29 (1975-76): 66-95.

Bremmer, Jan. *The Early Greek Concept of the Soul*. Princeton, N.J. Princeton University Press, 1983.

Burger, Ronna. *The Phaedo: A Platonic Labyrinth*. New Haven, Conn.: Yale University Press, 1984.

—. *Plato’s Phaedrus: A Defense of the Philosophic Art of Writing*. Birmingham: University of Alabama Press, 1980.

Burnet, John. “The Socratic Doctrine of the Soul.” *Proceedings of the British Academy* 7 (1916): 235-59.

Caston, Victor. “Epiphenomenalisms, Ancient and Modern.” *Philosophical Review* 106 (1997): 309-63.

—. "Toward a History of the Problem of Intentionality Among the Greeks." *Proceedings of the Boston Area Colloquium in Ancient Philosophy* 9 (1993): 213-45.

Claus, D. B. *Toward the Soul: An Inquiry into the Meaning of "Ψυχή" Before Plato*. New Haven, Conn.: Yale University Press, 1981.

Cooper, John M. "The Psychology of Justice in Plato." *American Philosophical Quarterly* 14/2 (1977): 151-57. Rpt. in John Cooper, *Reason and Emotion: Essays on Ancient Moral Psychology and Ethical Theory*. Princeton, N.J.: Princeton University Press, 1999.

Cornford, F. M. "The Division of the Soul." *Hibbert Journal* 28 (1929-30): 206-19.

—. "Psychology and Social Structure in the *Republic* of Plato." *Classical Quarterly* 6 (1912): 246-65.

—. *The Republic of Plato*. 1945. Oxford: Oxford University Press, 1971.

Crombie, I. M. *An Examination of Plato's Doctrines*. 2 vols. 1962. New York: Humanities Press, 1963.

Cross, R. C., and Woozley, A. D. *Plato's Republic: A Philosophical Commentary*. New York: St. Martin's Press, 1964.

Demos, Raphael. "Plato's Doctrine of the Psyche as a Self-Moving Motion." *Journal of the History of Philosophy* 6 (1978): 133-45.

Dodds, E. R. *The Greeks and the Irrational*. Berkeley: University of California Press, 1951.

—. "Plato and the Irrational." *Journal of Hellenic Studies* 65 (1945): 16-25.

—. "Plato's Views on the Nature of the Soul." In *Plato II: Ethics, Politics, and Philosophy of Art and Religion*. Ed. Gregory Vlastos. New York: Doubleday, 1971, 206-29.

Dorter, Kenneth. "Equality, Recollection, and Purification." *Phronesis* 17 (1972): 198-218.

—. "Plato's Image of Immortality." *Philosophical Quarterly* 26 (1976): 295-304.

Dybikowski, J. Critical notice of Thomas M. Robinson, *Plato's Psychology*. *Canadian Journal of Philosophy* 3 (1973): 131-42.

Elton, Matthew. "The Role of the Affinity Argument in the *Phaedo*." *Phronesis* 42 (1997): 313-16.

Ferrari, G. R. F. *Listening to the Cicadas: A Study of Plato's Phaedrus*. Cambridge: Cambridge University Press, 1987.

—. "The Struggle in the Soul: Plato, *Phaedrus* 253c7-255a1." *Ancient Philosophy* 5 (1985): 1-10.

Fine, Gail. "Separation," *Oxford Studies in Ancient Philosophy* 2. Oxford: Oxford University Press, 1984, 31-87.

—. "Immanence," *Oxford Studies in Ancient Philosophy* 4. Oxford: Oxford University Press, 1986, 71-97.

Gallos, Leon. "The State-Soul Analogy in Plato's Argument that Justice Pays." *The Journal of the History of Philosophy* 12/3 (1974): 285-93.

Gallop, David. *Plato: Phaedo*. Oxford: Clarendon [1975], 1988.

Gaye, R. K. *The Platonic Conception of Immortality, and its Connexion with the Theory of Ideas*. Cambridge: Cambridge University Press, 1904.

Gerson, Lloyd P. "A Note on Tripartition and Immortality in Plato." *Apeiron* 20 (1987): 81-96.

—. "Platonic Dualism." *Monist* 69 (1986): 351-69.

Gill, Christopher. *Personality in Greek Epic, Tragedy, and Philosophy: The Self in Dialogue*. Oxford: Clarendon, 1996.

—. "Plato and the Education of Desire." *Archiv für Geschichte der Philosophie* 67 (1985): 1-26.

Gottschalk, H. B. "Soul as Harmonia." *Phronesis* 16 (1971): 179-98.

Greco, Anna. "Plato's Cyclical Argument for the Immortality of the Soul." *Archiv für Geschichte der Philosophie* 78 (1996): 225-52.

Griswold, Charles L., Jr. *Self-Knowledge in Plato's Phaedrus*. University Park: Penn State Press, 1986.

Grote, G. *Plato and the Other Companions of Sokrates*. Vols. I-III. London: J. Murray, 1867.

Grube, G. M. A. *Plato's Thought*. Boston: Beacon Press, 1958.

Guthrie, W. K. C. *A History of Greek Philosophy*. Vol. 4. Cambridge: Cambridge University Press, 1987.

—. *Orpheus and Greek Religion: A Study of the Orphic Movement*. 1952. Princeton, N.J. Princeton University Press, 1993.

—. "Plato's Views on the Nature of the Soul." *Recherches sur la Tradition Platonicienne, Entretiens*. Tome III. *Pour l'Étude de l'Antiquité Classique*. Vandoeuvres-Genève: Fondation Hardt: 1957. Rpt. in *Plato II: Ethics, Politics, and Philosophy of Art and Religion*. Ed Gregory Vlastos. New York: Anchor Books, 1971, 230-41.

Hackforth, R. "The Modification of Plan in Plato's *Republic*." *Classical Quarterly* 7 (1913): 265-72.

—. "Immortality in Plato's *Symposium*." *Classical Review* 64 (1950): 43-45.

—. *Plato's Phaedo*. Cambridge: Cambridge University Press, 1972.

—. *Plato's Phaedrus*. Cambridge: Cambridge University Press, 1985.

—. "Plato's Theism." *Classical Quarterly* 30 (1936): 4-9.

Hall, Robert W. *Plato and the Individual*. The Hague: Martinus Nijhoff, 1963.

—. "Plato's Political Analogy: Fallacy or Analogy?" *The Journal of the History of Philosophy* 12/4 (1974): 419-35.

—. "Ψυχή as Differentiated Unity in the Philosophy of Plato." *Phronesis* 8 (1963): 63-82.

Halliwell, Stephen. *Plato: Republic 10*. Warminster: Aris & Phillips, 1988.

Hardie, W. F. R. *A Study in Plato*. Oxford: Clarendon, 1936.

Harrison, E. L. "Notes on Homeric Psychology." *Phoenix* 14 (1960): 63-80.

Hartman, Edwin. "Predication and Immortality in Plato's *Phaedo*." *Archiv für Geschichte der Philosophie* 54 (1972): 215-28.

Heath, M. "The Unity of Plato's *Phaedrus*." *Oxford Studies in Ancient Philosophy* 7, 151-73.

—. "The Unity of the *Phaedrus*: A Postscript." *Oxford Studies in Ancient Philosophy* 7, 189-91.

Hicken, W. F. "Phaedo 93a11-94b3." *Classical Quarterly* 48 (1954): 16-22.

Irwin, T. H. "The Parts of the Soul and the Cardinal Virtues." *Politeia*. Ed. O. Höffe. Berlin: Akademie Verlag, 1997.

—. *Plato's Ethics*. Oxford: Oxford University Press, 1995.

—. *Plato's Moral Theory*. Oxford: Oxford University Press, 1977.

Joseph, H. W. B. "Plato's *Republic*: The Nature of the Soul." In *Essays in Ancient and Modern Philosophy*. Oxford: Oxford University Press, 1935.

Kelly, John Clardy. "Virtue and Inwardness in Plato's *Republic*." *Ancient Philosophy* 9 (1989): 189-205.

Kenny, Anthony. "Mental Health in Plato's *Republic*." *Proceedings of the British Academy* 60 (1969): 229-53.

Keyt, David. "The Fallacies in *Phaedo* 102a-107b." *Phronesis* 8 (1963): 167-72.

Klagge, James C., and Nicolas D. Smith, eds. *Methods of Interpreting Plato and his Dialogues*. *Oxford Studies in Ancient Philosophy* suppl. vol., 1992.

Kraut, Richard. "Introduction to the Study of Plato." In *The Cambridge Companion to Plato*. Ed. Richard Kraut. Cambridge: Cambridge University Press, 1992, 1-50.

—. "Reason and Justice in Plato's *Republic*." *Exegesis and Argument*. Eds. E. N. Lee, A. P. D. Mourelatos, and R. M. Rorty. Assen: Van Gorcum, 1973, 207-24.

Lesses, Glenn. "Weakness, Reason and the Divided Soul in Plato's *Republic*." *History of Philosophy Quarterly* 4 (1987): 147-61.

Lovibond, Sabina. "Plato's Theory of Mind." In *Companions to ancient thought 2: Psychology*. Ed. Stephen Everson. Cambridge: Cambridge University Press, 1991, 35-55.

Matthews, Gareth B. "Commentary on Caston." *Proceedings of the Boston Area Colloquium in Ancient Philosophy* 9 (1993): 246-60.

McGibbon, D. D. "The Fall of the Soul in Plato's *Phaedrus*." *Classical Quarterly* NS 14 (1964): 56-63.

Miller, Mitchell. "Platonic Provocations: Reflections on the Soul and the Good in the *Republic*." In *Platonic Investigations*. Ed. D. J. O'Meara. Washington, D.C.: The Catholic University Press, 1985: 163-93.

Mills, K. W. "Plato's *Phaedo* 74b7-c6, part 1." *Phronesis* 2 (1957): 128-47.

—. "Plato's *Phaedo* 74b7-c6, part 2." *Phronesis* 3 (1958): 40-58.

Mohr, Richard. "The Sources of Evil Problem and the Principle of Motion Doctrine." *Apeiron* 14 (1980): 41-56.

Moline, Jon. "Plato on the Complexity of the Psyche." *Archiv für Geschichte der Philosophie* 60 (1978): 1-26.

—. *Plato's Theory of Understanding*. Madison: University of Wisconsin Press, 1981.

Moravcsik, J. M. E. "Understanding and the Emotions." *Dialectica* 36 (1982): 207-24.

Mulgan, R. G. "Individual and Collective Virtues in the *Republic*." *Phronesis* 13 (1968): 84-87.

Murphy, N. R. *The Interpretation of Plato's Republic*. Oxford: Oxford University Press, 1951.

Nehamas, Alexander. "Plato on Imitation and Poetry in *Republic* 10." In *Plato on Beauty, Wisdom and the Arts*. Eds. Julius Moravcsik and Philip Temko. Totowa, N. J.: Rowman & Littlefield, 1982.

—. "Predication and Forms of Opposites in the *Phaedo*." *Review of Metaphysics* 26 (1973): 461-94.

Nettleship, Richard Lewis. *Lectures on the Republic of Plato*. New York: St. Martin's Press (1902), 1961.

Neu, J. "Plato's Analogy of State and Individual." *Philosophy* 24 (1971): 238-54.

Nussbaum, Martha C. *The Fragility of Goodness*. Cambridge: Cambridge University Press, 1986.

O'Brien, D. "The Last Argument of Plato's *Phaedo* I." *Classical Quarterly* 17 (1967): 198-231.

—. "The Last Argument of Plato's *Phaedo* II." *Classical Quarterly* 18 (1968): 95-106.

Olszewsky, Thomas M. "On the Relations of Soul to Body in Plato and Aristotle." *Journal of the History of Philosophy* 14 (1976): 391-404.

Ostenfeld, Erik. N. *Ancient Greek Psychology and the Modern Mind-Body Debate*. Aarhus: Aarhus University Press, 1987.

—. *Forms, Matter and Mind*. The Hague: Martinus Nijhoff, 1982.

Pangle, Thomas. *The Laws of Plato*. New York: Basic Books, 1980.

Patterson, R. L. *Plato on Immortality*. University Park: Penn State Press, 1965.

Penner, Terry. "Plato and Davidson: Parts of the Soul and Weakness of Will." *Canadian Journal of Philosophy* suppl. vol. 16 (1990): 35-74.

—. "Thought and Desire in Plato." In *Plato II: Ethics, Politics, and Philosophy of Art and Religion*. Ed. Gregory Vlastos. New York: Anchor Books, 1971, 96-118.

Rees, D. A. "Bipartition of the Soul in the Early Academy." *Journal of Hellenic Studies* 77 (1957): 112-18.

Reeve, C. D. C. *Philosopher-Kings: The Argument of Plato's Republic*. Princeton, N.J.: Princeton University Press, 1988.

Robinson, James V. "The Tripartite Soul in the *Timaeus*." *Phronesis* 35 (1990): 103-10.

Robinson, Richard. "Plato's Separation of Reason and Desire." *Phronesis* 16 (1971): 38-48.

Robinson, T. M. *Plato's Psychology*. 2nd. ed. Toronto: Toronto University Press, 1995.

—. "Soul and Immortality in *Republic* X." *Phronesis* 12 (1967): 147-51.

Rohde, Erwin. *Psyche*. Trans. W. B. Hillis. 8th ed. New York: Books for Libraries Press, 1972.

Rowe, C. J. "The Argument and Structure of Plato's *Phaedrus*." *Proceedings of the Cambridge Philological Society* 32 (1986): 106-25.

\_\_\_\_\_. *Plato: Phaedo*. Cambridge: Cambridge University Press, 1993.

\_\_\_\_\_. *Plato: Phaedrus*. Warminster: Aris & Phillips, 1986.

\_\_\_\_\_. "The Unity of the *Phaedrus*: A Reply to Malcolm Heath." *Oxford Studies in Ancient Philosophy* 7, 175-88.

Saunders, T. J. "The structure of the soul and the state in Plato's *Laws*." *Eranos* 60 (1962): 37-55.

Scaltsas, Theodore. "Soul as Attunement: An Analogy or a Model?" In *Greek Studies in the Philosophy and History of Science*. Ed. Pantelis Nicolaopoulos. Dordrecht: Kluwer, 1990.

Scarraw, D. S. "Phaedo 106a-106e." *Philosophical Review* 70 (1961): 245-52.

Schipper, E. W. "Souls, Forms and False Statements in the *Sophist*." *Philosophical Quarterly* 15 (1965): 240-42.

Shorey, Paul. *The Unity of Plato's Thought*. New York: Archon Books [1903], 1968.

Skemp, J. B. *The Theory of Motion in Plato's Later Dialogues*. Amsterdam: Hakkert, 1967.

Solmsen, Friedrich. "Plato and the Concept of the Soul (*Psyche*): Some Historical Perspectives." *Journal of the History of Ideas* 64 (1983): 355-67.

Stalley, R. F. "Plato's Argument for the Division of the Reasoning and the Appetitive Elements within the Soul." *Phronesis* 20 (1975): 110-28.

Stewart, J. A. *The Myths of Plato*. New York: Macmillan, 1905.

Stocks, J. L. "Plato and the Tripartite Soul." *Mind* 24 (1915): 207-21.

Taylor, A. E. *Plato: The Man and His Work*. New York: Meridian Books, 1956.

Tiles, J. F. "The Combat of Reason and Passion." *Philosophy* 52 (1977): 321-30.

Tsirpanlis, E. C. "The Immortality of the Soul in *Phaedo* and *Symposium*." *Platon* 17 (1965): 224-34.

Vlastos, Gregory. "Justice and Happiness in the *Republic*." In *Plato II: Ethics, Politics, and Philosophy of Art and Religion*. Ed. Gregory Vlastos. New York: Anchor Books, 1971, 66-95.

\_\_\_\_\_. "Justice and Psychic Harmony in the *Republic*." *Journal of Philosophy* 66 (1969): 505-21.

\_\_\_\_\_. *Socrates, Ironist and Moral Philosopher*. Ithaca, N.Y.: Cornell University Press, 1991.

\_\_\_\_\_. "The Theory of Social Justice in the *Polis* in Plato's *Republic*." In *Interpretations of Plato*. Ed. Helen North. Leiden: E. J. Brill, 1977, 1-40.

Wagner, Ellen. "Unity and Development in Plato's Psychology." *The Journal of Neoplatonic Studies* 5 (1997): 81-95.

White, David. A. *Myth and Metaphysics in Plato's Phaedo*. London: Associated University Presses, 1989.

\_\_\_\_\_. *Rhetoric and Reality in Plato's *Phaedrus**. Albany: State University of New York Press, 1992.

White, Nicholas. *A Companion to Plato's Republic*. Indianapolis, Ind.: Hackett, 1979.

Wilford, F. A. "The Status of Reason in Plato's Psychology." *Phronesis* 4 (1959): 54-58.

Wilson, J. R. S. "The Argument of *Republic* IV." *The Philosophical Quarterly* 26 (1976): 111-24.

Wolfe, J. "A Note on Plato's 'Cyclical Argument' in the *Phaedo*." *Dialogue* 5 (1966): 237-38.

Woods, Michael. "Plato's Division of the Soul." *Proceeding of the British Academy* 73 (1987): 23-48.

## Index

Ackrill, J. L., 325  
Adam, James, 297, 309–11  
Adeimantus, 115, 122, 123, 126  
affections (*pathē, πάθη*), 348; in *Laws*, 215–17, 221, 244  
Affinity Argument for immortality, 6, 8, 9, 13, 14, 19, 69–88, 141–42, 145, 147, 148, 149, 257, 259, 290, 293, 294  
*aitia*, 82–83  
*akrasia*, 20; models of akratic action, 205–9, 213; ‘hear’ akratic action, 218; Plato’s acceptance of, 137–40; and Reciprocity of Virtues, 223–226; Socratic view of, 137–40; 151, 203–26; ‘strict’ akratic action, 214–18  
Alcmaeon, 15  
analogy of soul and state, 19, 116, 122, 127, 131, 157–58, 170, 176–88  
*Anamnesis* Argument. *See* Recollection Argument  
Annas, Julia, 24, 123–25, 297, 307, 311, 312, 315, 316, 317  
appetitive man, 181–82  
appetitive part of soul, 10, 12, 19, 94, 95, 180–82, 219–22; in alliance with spirit, 98; desires of, 120, 207; how persuaded, 210–12; pleasures of, 100–102; unifying principle of, 99–102, 232nn19–20  
Archer-Hind, R. D., 11, 60  
Aristotle, 52, 55, 56, 75, 93, 248; and *akrasia*, 218; and *athanasia*, 278n5; and kinds of desires, 109n2; on self-motion, 356n8  
*athanatos* (ἀθάνατος), 8, 15; ambiguity of term, 285  
Barnes, Jonathan, 263–77  
Bett, Richard, 21–22  
Bluck, R. S., 58, 60  
Bobonich, Christopher, 20  
body, as ensouled, 42–43  
Bostock, David, 20  
Brandwood, Leonard, 24  
Brown, Eric A., 21  
Burnet, John, 60  
Cebes, 14, 73, 74, 78, 82, 83, 264–65, 272, 274–76, 281, 282, 286, 290  
cities in *Republic*, 176–88; “original” city, 122–24; “swollen” city, 124. *See also* Kallipolis  
conflict in soul, 10–11, 95; 97–98, 113n19, 166–67, 209, 212; and poetry, 188–89  
consciousness, 26n5  
Cooper, John, 19  
Cornford, F. M., 3, 12  
Crombie, I. M., 12  
Cyclical Argument for immortality, 7, 13, 14, 20, 21, 258, 263–77, 293; and ‘law of opposites,’ 282–83, 295n4; objections to, 274–77; supplanted by Final Argument, 277  
death and dying, 241–42, 246, 255; and Socrates, 249

*de dicto, de re* necessity, 149–50  
 democratic city, 185–86, 192  
 democratic man, 183–84, 191, 195n28  
 desire, desiring part of soul. *See*  
     appetitive part of soul;  
     *epithumētikon*  
*dkaios* (δίκαιος), 157, 159–63  
*dkaiosune* (δίκαιοσύνη), 157, 159–62  
 Dodds, E. R., 4  
 Dualism, “Platonic,” 3, 69; substance, 18, 75, supervenient, 69, 76; epiphenomenalism, 76

Echecrates, 56  
 emotions. *See* affections  
 Empedocles, 58  
*epithumētikon* (ἐπιθυμητικόν),  
     *epithumia* (ἐπιθυμία), 10, 11, 16, 38, 159, 167, 179, 194, 204, 215, 257; and worker class of republic, 163–67; desires of *epithumētikon*, 138. *See also* appetitive part of soul  
 externalization in *polis–psychē* relation, 174–76, 177–78, 187–88

Final Argument for immortality, 6, 13, 14, 18, 35–49, 81, 260, 281–94; and ‘law of opposites,’ 287; as supplanting Cyclical Argument, 277; objections to, 284–94

folk psychology, 2

Forms, 1, 5, 18, 38–39, 48n8, 70, 71; and soul, 132, 141–42, 245–46, 288–90; 291

forms, immanent. *See* immanent forms

Ferrari, G. R. F., 12

Frede, Dorothea, 21

Freud, Sigmund, 2, 201nn144, 149

Gallop, David, 20, 35, 58, 59, 63, 64, 80, 81

*gignesthai* (γίγνεσθαι), 270–73, 279n17

Glaucon, 21, 115, 116, 122, 123, 125–27, 128, 129, 298, 299, 301, 302–11, 314

Glaucus, 128, 143, 144, 145, 148, 294  
 gods, appetites of, 189–90

Guthrie, W. K. C., 128

Hackforth, R., 60

Hades, 4, 22, 247, 265, 266, 267, 268, 270, 274, 275, 281

Halliwell, Stephen, 297

*harmonia*, 51–67, 259, 290; causal dependency on body, 53; control of its elements, 59; definitions of, 52–57, 75–76; degrees of, 79–80; types of, 79–80; as Form of soul, 74; as immaterial, 72

*harmonia* thesis (*Phaedo* 85e–86a), 2, 6, 8, 18, 19, 27n22, 51–67; arguments against, 51, 58–66, 79–82, 259–60; epiphenomenal interpretation of, 76; interpretations of, 52; 69–88, 84n2; as nonmaterialist, 73–76, 85n10; as supervenient, 76–79

Hartman, Edwin, 291–92, 293

Hicken, W. F., 58, 60, 65

Hobbes, Thomas, 93

Homer, 4, 102, 166, 170

homunculi, 213

*Iliad*, 4

immanent characters, 35, 36, 37, 40–44, 47nn5–6, 291

immanent forms, 6, 8, 18, 38, 40, 85–86n16, 289, 291

Immortality, arguments for. *See*

    Affinity Argument, Cyclical Argument, Final Argument, Recollection Argument, *Phaedrus*, *Republic*  
     imperishability of soul, 264

internalization in *polis–psychē* relation, 171–73, 176, 187–88, 195–96n36, 46

Irwin, T. H., 95, 153n1

justice, as doing one’s own, 19, 126–27, 129–30, 160; as harmony between inside and outside of *psychē* and *polis*, 187–88; images of, 125–26; of philosopher, 252; as the same in soul and state, 92

*Kallipolis*, 116–17, 122–26, 129, 130, 132, 303

Keyt, David, 18, 36–44, 46, 47

*kinesis* (κίνησις), 339–41

Kripke, Saul, 150–51

Lear, Jonathan, 20

Leontius, 98, 99, 100, 101, 102–4, 106, 166, 213

*logistikon* (λογιστικόν), 10, 11, 94, 257; desires of, 138; 159, 160–61; 163–67, 179, 180, 204; and worker (appetitive) class, 164, 179. *See also* reason

McLaughlin, Brian, 76, 77

*Medea*, 4

Melissus, 316–17

*menos*, 4, 27n11

mereology, 17, 19, 144, 148

*meros, merē*, (μέρος), 11, 156n24, 221, 236n49

Miller, Mitchell, 153n1

Nails, Debra, 24

Nettleship, R. L., 312

*noos*, 4, 27n11

Nussbaum, Martha, 12, 360n33

O’Brien, D., 264

*Odyssey*, Odysseus, 4, 102, 103, 104, 106, 107, 267

oligarchic man, soul of, 180–1, 185, 191

Olympiodorus, 264

Orphic–Pythagorean beliefs, 3, 5, 22, 27nn16, 17, 248

parts of soul: abstract parts, 148; agent-like parts, 17, 20, 203, 204, 208, 212–14, 216–18, 220–21, 222, 234n40; aggregative parts, 146–48; as psychological

determinants, 93–94; conceptual parts, 146–48; contingent parts, 145; and cultural influence, 186–87; definition of ‘part’, 147; essential parts, 149–50; functional parts, 2, 147; organic parts, 146–48; proper parts, 145–46; spatial parts, 148

pathology in the city, 171–72, 183

pathology in the soul, 183, 187–88, 191, 312

perception, 243–44, 245

*Phaedrus*, argument for immortality in, 15, 17, 323–28, 335–53, 354n3, 354–55n5

Philolaus, 54, 56, 57, 73, materialist conception of soul, 78

philosopher–king, 174–75, 225  
*phrenes*, 4, 27n11

poetry, effect on soul, 188–93; and intrapsychic conflict, 188–89

Predication, Pauline, 291, 296n20

Principle of Opposites: (Part–Generating Principle; *Rep.* 438b8), 9, 27–28n23, 117–21, 145, 148, 149, 205–6, 211–12; not in *Laws*, 219; and intensional/extensional identity, 119; whether logical or empirical, 146–49

Proclus, 314–17

properties, accidental, 283–85, 292

properties, essential, 283–85, 286, 291

*psuchē*, 4, 140, 171–77; falsehood in, 175; meanings of word for Greeks, 242–44; use of term, 352, 359–60nn23–24

psychology, Plato’s, contradictions in, 3; developmental view of, 138–41

purification of soul (*katharsis*), 5, 10, 71, 246, 248, 274

Pythagoras, 56

reason (reasoning part of soul), 3, 11, 16, 19, 94, 96, 220–21; and desire for good, 97, 207–8; how persuaded, 211–12

Receptacle, The, 326, 349

Recollection Argument, 13, 14, 20, 258–59, 266, 268, 290, 293

Reeve, C. D. C., 123–25

reincarnation, 16, 248, 252, 253–54, 256–7, 298, 314–17

*Republic*, argument for immortality in (*Rep.* X, Special Vice Argument), 13, 15–16, 142–44, 148, 154n8,

257, 297–317, 318n7; as ethical argument, 312; objections to, 307–14  
 Robinson, T. M., 21  
 Sedley, David, 82–83  
 separation, of soul and body, 56, 241, 242, 245, 246, 247, 264–65, 294, 305–6, 308–11, 321n23  
 Shields, Christopher, 19, 77  
 Simmias, 6, 8, 18, 19, 51, 53, 54, 56, 58, 64, 69, 70–74, 76–79, 81–83, 259, 260, 276, 281–83, 289, 290  
 Smith, Nicholas D., 19  
*sôphrosunê*, 161  
 soul, activities of, 242–44; and consciousness, 254–56; and education, 170–72; and the person, 253–58; as alive, 42–46, 48n7; as always in motion, 323–25; as *archê*, 324–25, 338–39; as breath, 5; as chariot, 22, 166–67, 219, 257, 347–48; as constant, 70; as Form, 18; as essentially composite, 149; as form of a body, 46; as immaterial, 3, 253; as imperishable, 35–36; as indissoluble, 7; as nonphysical, 54–55; as principle of life, 293; as prisoner of body, 56; as seat of violent emotions, 5; as self-mover, 15, 17, 21, 323–28, 338–40, 341, 344–45; as simple, 9, 17, 111–12, 128, 138, 141–42, 144, 151; as supervenient, 69; brings life to body, 270–73, 279n15, 284, 292; causal dependency on body, 52; causal power of, 209; conceptions of in *Phaedo*, 5; good and bad souls, 79–81; in *Timaeus*, 258; instability of, 7, 11; modularity of, 2; natural evil of, 300–8, 319–20nn11–17; of dead, 248; ordering of, 59–63; resemblance to Forms, 6; substance, 7, 19, 35, 287–88, 290; true nature of, 128, 144  
 soul “stuff,” 7, 341–42, 358–59n22

spirit (spirited part of soul; *see also thumos*), 4, 10, 12, 19, 94, 95; in *Rep.* VIII–IX, 104–7; in *Laws*, 219; in alliance with appetite, 98  
 Stewart, J. A., 297, 308  
 Strato of Lampsacus, 285  
 supervenience, 2, 69–88; definition of, 76–77; substance supervenience, 77–78; problem of dualism in, 86–87n23; requirements for, 87n29  
 Taylor, C. C. W., 18, 80, 81  
 temperance (*σωφροσύνη*), in *polis*, 182–83; in soul, 249–52  
 Thrasymachus, 115  
*thumos, thumoeides* (*θύμος, θυμοειδές*), 4, 10, 11, 16, 27nn11, 15, 113nn18, 19, 116, 159, 179, 204, 215, 221, 257; and military class of republic, 165–66; soul ruled by, 104–5; examples of in *Rep.* IV argument, 106–7; nature of, 166–67; origin of *thumos*—desires, 107–8; unifying principle of, 102–7  
 tripartition, argument for (*Rep.* IV), 12, 16, 19, 108, 116, 127–28, 130, 139, 140, 141, 144–49, 151, 348, 350; consistency with *Phaedo* and *Rep.* X, 241–51, 346–49; objections to, 212–214  
 tyrant, soul of, 191–92  
 unjust souls, types of in *Rep.*, 101, 104–5, 112n13  
 virtue, human, in *Phaedo*, 250–52; in *Republic*, 91–92  
 Virtues, Reciprocity of, 223–26  
 Vlastos, Gregory, 24, 225  
 Wagner, Ellen, 18  
 Williams, Bernard, 19, 176–79, 185–87, 199nn101, 103, 107  
 World Soul, 16, 323, 327, 362n36, 340–42, 350–51, 358n19

## Index Locorum

### I. Plato

*Phaedo*  
 61d6–9, 78  
 61d–e, 56  
 63b–c, 246  
 64c4–8, 241  
 64d4–7, 245  
 64d8–e6  
 65a, 5  
 65a9–b7, 245  
 65b–d, 5  
 66a2, 245  
 66b–67b, 5  
 66b7–c5, 242  
 67c–d, 5  
 67c8, 247  
 68c8–10, 249  
 68d, 249  
 68d12, 250  
 68e3, 249  
 69a6, 251  
 69a9–10, 252  
 69b3, 251  
 69b8–c3, 251  
 69e, 246  
 69e–72d, 13  
 69e5–70a7, 305  
 69e6–72e2, 20  
 70b2–4, 265  
 70b5–6, 265  
 70c4–5, 266, 275  
 70c5–6, 265  
 70c8–d2, 271, 272  
 70d–71a, 7

70d4-5, 274  
 71c-d, 286  
 71e2, 266  
 72a-d, 258  
 72e-77d, 14  
 73a1-3, 9  
 73a2, 264, 271  
 74b4-9, 70  
 74c5, 70  
 75c11-d1, 70  
 76a1-9, 70  
 76d6-9, 70  
 77b1-c5, 275  
 77b3-6, 305  
 77c1-5, 14  
 77d2-4, 242  
 77d5-e2, 275  
 77e1, 247  
 78b-80b, 70  
 78b4-84b8, 14, 141  
 78c-80c, 38  
 78c1-2, 7, 141  
 78c1-4, 141  
 78d10-e1, 70  
 79a-c, 7  
 79a1-2, 73  
 79a3, 9  
 79a6-7, 14, 70, 73  
 79b, 14  
 79b16, 70  
 79c5-9, 7  
 79c-e, 259  
 79d5, 70  
 79d8, 7  
 79e3-4, 14  
 79e4-5, 70  
 80a, 259  
 80a7, 76  
 80a8, 70  
 80a9, 14, 76  
 80b1, 9, 38, 71, 264  
 80b3-4, 71  
 80b9-10, 14, 71  
 80b10, 141, 144  
 80d, 268, 274  
 80d2, 71  
 80d8-e1, 305  
 81a, 246, 264  
 81a9, 71

81b, 5  
 81c4-6, 9, 247  
 81c11, 247  
 81e1, 247  
 82c5-8, 249  
 82e, 5  
 83d, 5  
 83d4, 9  
 84b2, 242  
 85b5-c1, 53  
 85d, 73  
 85e5-6, 53, 73  
 85e-86a, 6, 8, 18, 51, 53, 71, 72, 73,  
     259  
 86a1, 72  
 86a2, 71  
 86b1-2, 73  
 86b7-c1, 8  
 86b6-c2, 51, 71, 73  
 86b8-c1, 79  
 86b9, 55, 71  
 86c6-7, 54  
 87d-88b, 265  
 88a6, 242  
 88a10-b2, 265  
 88b6-7, 56  
 88c, 51  
 88d3-4, 56  
 91e-92e, 57  
 92a-95a, 6, 8, 18, 69, 78  
 92a6-95a2, 18  
 92a8, 84n7  
 92c-93a, 259  
 92d2, 54  
 92e4-93a10, 58, 59, 65  
 92e6-95a2, 18, 69, 78-80  
 93a1, 84n7  
 93a4-5, 59, 65  
 93a6-7, 59, 65  
 93a8-9, 59, 65  
 93a11-12, 58, 59, 65, 79, 80  
 93a11-c10, 79  
 93a14-b2, 59, 65  
 93a-94b, 259  
 93b1, 79  
 93b4-6, 59, 65, 79  
 93b8-c2, 59, 65, 79, 80, 267  
 93c310, 59, 65  
 93d1-5, 59, 66, 78

93d6-8, 61, 66, 79  
 93d9-11, 61, 66, 79  
 93d12-e2, 61, 66, 79  
 93e4-5, 61, 66, 79  
 93e7-10, 79  
 93e7-94a10, 61, 66  
 94a12-b3, 58, 79  
 94b-c, 76  
 94b4-5, 267  
 94b-95a, 259  
 94b4-c1, 64, 66, 76  
 94c3-6, 64, 66  
 94c9-d5, 76  
 94e8-95a2, 61, 66  
 95a2, 58  
 95c-102a, 260  
 95c-99d, 294  
 95d5, 265  
 95e-96a, 282  
 100b3-4, 295n8  
 100c4-5, 296n10  
 100d4, 38  
 100d8, 38  
 102a-107a, 6, 21, 74  
 102a10-107b10, 18, 35, 36, 38, 39, 40,  
     42, 44, 74, 81  
 102b-105c, 260  
 102c6, 8  
 102d-e, 268, 282  
 102d5, 37, 39  
 102d5-103c8, 81  
 102d6, 296n10  
 102e5-7, 8  
 103a1, 282, 283  
 103a4-b9, 39  
 103a4-c4, 286  
 103b, 290, 295n6  
 103b5, 74  
 103b-c, 291  
 103b8-9, 292  
 103d-e, 268  
 103d5-8, 40  
 103e, 292  
 103e2-5, 40, 42  
 104c7-9, 7  
 104d1, 8  
 104d1-3, 284-85, 288, 289, 290  
 104d5-7, 37, 290  
 104d5-e10, 288

*Phaedrus*  
 245c-e, 264  
 245c-246a, 15, 17, 21  
 245c1-2, 323  
 245c1-8, 15  
 245c5, 324, 336, 337  
 245c7-8, 336, 338  
 245d1, 338, 339  
 245d1-6, 15  
 245d2-3, 340  
 245d3-4, 340

245d4-6, 340  
 245d6-e2, 15  
 245d6-7, 339  
 245d7-8, 339  
 245e2-3, 336, 344  
 245e2-246a2, 15, 357n24  
 245e3-246a1, 336  
 245e4-6, 344  
 245e5, 337  
 246a, 11  
 246a-247c, 6, 351, 358n20  
 246a1-2, 336  
 246a6, 336  
 246b6, 351, 359n23  
 246e3, 347  
 247c-248e, 348  
 247d3, 348  
 248c2-e3, 356n20  
 248e-249d, 15  
 249b-c, 348  
 254c, 167  
  
*Protagoras*  
 352b-358e, 140  
 358e, 137  
  
*Republic*  
 352b-c, 138  
 353d4-6, 130  
 353d9, 130  
 353d11-e11, 130  
 354a-c, 115  
 354a12-b5, 115  
 354b5-c1, 115  
 358b, 171  
 358e, 137  
 368c7-d7, 115, 142  
 368d, 157  
 369a5-7, 124  
 369c6-7, 123  
 369c9-10, 123  
 370a-b, 126  
 371e, 122  
 372a3-4, 122, 124  
 372d4-5, 122  
 372e2-6, 126  
 372e6-7, 122  
 373b, 189  
 373d-e, 189

374a-e, 126  
 375a11-12, 104  
 375e10, 104  
 377a-b, 171  
 377e, 190  
 377e-378d, 189  
 378b-c, 190, 191  
 378d-e, 171, 190  
 380b-c, 171  
 382, 191  
 382b-c, 175  
 387, 172  
 393b-d, 188  
 395c-d, 172  
 395c-396e, 188  
 399e, 189  
 400d-e, 175, 189  
 400e-401a, 175  
 401b-c, 173  
 401d-e, 173  
 411, 188  
 411e-412a, 172  
 412e-413b, 190  
 413a4-5, 207  
 419a, 158  
 421c, 180  
 421d-422a, 124  
 422e, 172  
 424c-d, 173  
 429b-430a, 183  
 431b-d, 161  
 431d-e, 182  
 431e-432a, 161, 182  
 432d-433a, 126  
 433c-d, 160  
 433e-434a, 182  
 433e3-8, 125, 161  
 434b9-10, 125  
 434b9-d1, 116, 142  
 434e, 157  
 435a-b, 178  
 435b, 117,  
 435c, 174  
 435b-c, 93, 157  
 435e, 158, 174, 177  
 435e-436a, 109n4, 174  
 436a11, 98  
 436b, 9  
 436b-c, 95, 96

436b8, 145  
 436b8-9, 205  
 436b8-c1, 117  
 436b-441a, 141, 145  
 436c5-6, 117  
 436e1, 148  
 437b-c, 118, 145  
 437b1-5, 118, 145, 206  
 437b1-d1, 205  
 437d2-3, 98  
 437d2-439b1, 211  
 437d8-10, 145  
 437d11-e2, 112n11  
 438a, 138  
 439a1-e1, 118  
 439a-e, 118, 150  
 439a9-c8, 206  
 439b1, 95  
 439b8-11, 96  
 439c1-2, 145  
 439c5-6, 209  
 439c7, 210  
 439d1-2, 99, 120  
 439d6, 11, 98  
 439e-440a, 10  
 439e-440e, 166  
 439e2, 11  
 440a-c, 98, 100  
 440a1, 98  
 440a2-3, 100  
 440b2-4, 113n19  
 440b4-7, 113n19  
 440c7-9, 113n19  
 440d, 166  
 440e-441a, 166  
 441a-b166  
 441a2-3, 113n19  
 441a7-9, 101  
 441a-b, 102, 106, 166  
 441d-e, 187  
 441e4-5 94, 95  
 441b2-3, 102  
 441b4-c2, 102  
 442b5-c3, 103  
 442b11, 11  
 442c5-8, 94, 95, 207, 210  
 443b-d, 126  
 443c-444a2, 131  
 443c9-d7, 128

443d7, 155n23  
 446e-467a, 172  
 478e-479b, 125  
 484c, 174  
 500c, 174  
 505d5-e4, 207  
 505e-506a, 140  
 508c2, 112n14  
 544d, 162, 174  
 546a1, 303  
 548e4-549a7, 104  
 549a9-b2, 104  
 550b4-6, 113n18  
 550b7, 104  
 551b, 163  
 551d, 172  
 553c-554e, 172  
 553c4-d7, 113n18  
 553d4-7, 105, 113n19  
 554-555, 184  
 554b-e, 172, 180  
 554b7-c2, 113n19  
 554c11-d3, 220  
 554c11-e5, 211  
 554d2, 210  
 554d6-e5, 109n4, 184  
 554d9-e1, 96, 110n7  
 556b-c, 184  
 557a, 163  
 557c, 162, 185  
 557d-558a, 185  
 558-562, 184  
 558d-559c, 180  
 559d-561e, 112n13, 184  
 559e, 185, 189  
 560, 185  
 561b-d, 162  
 561b2-c3, 98  
 561c-d, 99, 185  
 564d, 162  
 565d-566b, 190  
 566a-b, 192  
 567a, 162  
 567b-c, 191  
 568a, 162, 190  
 568b, 192  
 569b, 192  
 571b-d, 189  
 571c, 192

572b, 190  
 572c-573c, 191  
 573b, 190  
 574a-c, 191  
 575a-c, 162  
 577a, 192  
 577c, 162, 192  
 580d-581b, 213  
 580d-581e, 182  
 580d-588a, 360n27  
 580d7-8, 94, 360n27  
 580e, 100  
 580e5-581a1, 112n12  
 581, 182  
 581a9-10, 103  
 581b5-6, 94  
 581c, 167, 182, 183  
 586c7-d2, 361n27  
 588c-e, 10  
 589a6-b6, 211  
 590e, 173  
 592b, 174-75  
 598b, 188  
 601b, 188  
 602b, 189  
 603a, 188  
 603d6-7, 128  
 604d, 189  
 604e-605d, 189  
 605, 188  
 605b-c, 188  
 605d-e, 97  
 608b6, 310n2  
 608c-611a, 15, 143  
 608c1-2, 318n2  
 608c3, 298  
 608c9-d1, 264  
 608d3-12, 298  
 608d13-e5, 299  
 608e3-5, 318n7  
 608e6-609a5, 299-300  
 609a6-8, 300  
 609a8, 154n8  
 609a9-b3, 301  
 609a10-b1, 302  
 609b4-7, 303  
 609c2-5, 306  
 609c2-d3, 304  
 609d4-8, 304  
 609d6-7, 308  
 609d9-610e9, 305  
 610a5, 320n17  
 610a8-9, 305  
 610d3-4, 306  
 610d7-e4, 310  
 610e10-611a2, 305, 306  
 611a, 294  
 611a1, 154n8  
 611a2, 264  
 611a4-9, 314  
 611b, 17, 128, 142, 143, 144  
 611b2-3, 346  
 611b5-6, 346  
 611b9-d8, 128  
 611b10-c1, 128  
 611b-612a, 257  
 611e2, 142  
 611e4-5, 128  
 612a, 144  
 612a3-5, 128  
 612a5-6, 128  
 612a8, 298  
 613c6, 318n2  
  
*Sophist*  
 248a4-249d4, 345  
  
*Timaeus*  
 35a1, 348  
 41d, 356n19  
 41d-44c, 16,  
 42a-b, 362n35  
 42a3-4, 362n35  
 49a, 21  
 52e-53c, 326  
 53a2, 326  
 53a8, 327  
 53c4, 288  
 57c, 288  
 69c-70b, 6, 16, 348  
 69c7-8, 16  
 69c3-70a2, 219  
 69c8, 350  
 71b5, 221  
 71d1-2, 221

**II. Other Ancient Authors**

Aristotle  
*De Anima*  
 407b27-8, 56  
 407b27-408a30, 8  
 408a5-21, 75  
 408a20-1, 55  
 408a30-1, 55  
 56, 75  
  
*De Mundo*  
 396b, 53  
  
*Nicomachean Ethics*  
 1095b, 167  
  
*Physics*  
 201a10-16, 340

Empedocles  
 DK 31B, 52, 58  
 DK 22B, 52,  
  
 Euripides  
*Medea* 1021-80, 4  
  
 Macrobius  
 DK 44A, 56  
  
 Parmenides  
 DK B8.26-8, 340  
  
 Philolaus  
 DK 44B, 56-57

## About the Contributors

**Richard Bett** is Professor of Philosophy and Classics at Johns Hopkins University. He is the author of *Pyrrho, his Antecedents and his Legacy* (Oxford, 2000) and of a translation, with introduction and detailed commentary, of Sextus Empiricus' *Against the Ethicists* (Oxford, 1997). He has also published articles in *Phronesis*, *Oxford Studies in Ancient Philosophy*, *Apeiron*, and elsewhere.

**Christopher Bobonich** is Assistant Professor of Philosophy at Stanford University. He is the author of several articles in ancient philosophy and of *Utopia Recast: Plato's Later Ethical and Political Philosophy*, forthcoming from Oxford University Press.

**David Bostock** is Fellow and Tutor in Philosophy at Merton College, Oxford. His books include *Plato's Phaedo* (1986), *Plato's Theaetetus* (1988), *Aristotle's Metaphysics Z and H* (1994), and *Aristotle's Ethics* (2000).

**Eric A. Brown** is Assistant Professor of Philosophy at Washington University in St. Louis. He is currently completing a book called *Stoic Cosmopolitanism*. He has also published, among other works, "Justice and Compulsion for Plato's Philosopher-Rulers," *Ancient Philosophy* 20 (2000): 1-17 and "Hellenistic Cosmopolitanism" in *A Companion to Ancient Philosophy*, ed. Mary Louise Gill and Pierre Pellegrin (Oxford: Basil Blackwell, forthcoming).

**John M. Cooper** is Stuart Professor of Philosophy, Princeton University. He is author of *Reason and Human Good in Aristotle* (1975) and *Reason and Emotion: Essays on Ancient Moral Psychology and Ethical Theory* (1999), as well as editor (with J. Procopé) of *Seneca: Moral and Political Essays* (1995) and of *Plato: Complete Works* (1997).

**Dorothea Frede** studied philosophy and classics in Germany, where she obtained her Ph.D. at Goettingen University. From 1971-1991 she lived and taught in the United States at various colleges and universities (San Francisco State, Berkeley, Stanford, Princeton University, Rutgers University, and Swarthmore College). In 1991 she accepted a chair in ancient philosophy at Hamburg Uni-

versity. She has published widely in ancient philosophy, mostly on Plato and Aristotle, but also in Hellenistic and Later Greek and Roman philosophy. Among her works are a translation of the *Philebus* for Hackett and a commentary on the same dialogue in German (1997). She is currently working on a monograph for Cambridge University Press: *Cosmology, Rationality, and Politics: Toward a New Interpretation of Plato's Timaeus*.

**David Gallop** is Professor Emeritus, Department of Philosophy, at Trent University, Ontario, where he taught from 1969 to 1989. He has published numerous articles on philosophical and literary subjects. He has translated and edited *Plato's Phaedo* for the Clarendon Plato series (Oxford 1975), as well as *Euthyphro, Defence of Socrates, Crito and Phaedo* for World's Classics (1993, 1997). He has also published *Parmenides of Elea* (Toronto 1984) and *Aristotle on Sleep and Dreams* (Warminster 1996).

**Jonathan Lear** is John U. Nef Distinguished Service Professor at the Committee on Social Thought, The University of Chicago. He is the author of *Happiness, Death and the Remainder of Life*, *Open Minded: Working Out the Logic of the Soul, Love and Its Place in Nature: A Philosophical Interpretation of Freudian Psychoanalysis*, and *Aristotle: The Desire to Understand*. He did his graduate work in philosophy at Cambridge University and The Rockefeller University; he trained as a psychoanalyst at the Western New England Institute of Psychoanalysis. Before coming to the University of Chicago, Professor Lear taught at Cambridge University, where he was a Fellow of Clare College and at Yale University, where he was Kingman Brewster Professor of the Humanities.

**Thomas M. Robinson**, Professor of Philosophy at the University of Toronto, is the author of *Plato's Psychology* (Toronto: University of Toronto Press, 1970; second edition 1995), and has published widely on topics in Greek philosophy. From 1993 to 1995 he was President of the Society for Ancient Greek Philosophy and from 1995 to 1998 President of the International Plato Society. Since 1993 he has also been Honorary President of the International Association for Greek Philosophy.

**Christopher Shields** is Professor of Philosophy and Classics, University of Colorado at Boulder. He has published on Plato, Aristotle, and Stoicism. His book *Order in Multiplicity Homonymy in the Philosophy of Aristotle* was published in 1999 by Clarendon Press; forthcoming is a translation of and commentary on the *De Anima* in the Clarendon Aristotle Series.

**Nicholas D. Smith** received his Ph.D. from Stanford University in 1975 and is now James F. Miller Professor of Humanities at Lewis and Clark College in Portland, Oregon. He is the author, coauthor, editor, or coeditor of seventeen books and numerous articles on Aristophanes, Aristotle, Plato, and Socrates.

Recent books include *Plato's Socrates* (with T. C. Brickhouse, Oxford University Press, 1994), *The Philosophy of Socrates* (with T. C. Brickhouse, Westview Press, 1999), *Reason and Religion in Socratic Philosophy* (edited with Paul Woodruff, Oxford University Press, 2000), and *The Trial and Execution of Socrates* (edited with T. C. Brickhouse, Oxford University Press, 2001).

**C. C. W. Taylor** is a Fellow of Corpus Christi College, Oxford, and Reader in Philosophy at Oxford University. His publications include *Plato*, *Protagoras* (Oxford 1991), *Socrates* (Oxford 1998), *The Atomists: Leucippus and Democritus* (Toronto 1999), and (jointly with JCB Gosling) *The Greeks on Pleasure* (Oxford 1982). He is the editor of *Routledge History of Philosophy Vol I: From the Beginning to Plato* (London and New York: Routledge, 1997), and was the editor of *Oxford Studies in Ancient Philosophy* from 1993 to 1998.

**Ellen Wagner** is Associate Professor of Philosophy at the University of North Florida. She has published articles on Plato in the *Journal of Neoplatonic Studies*, *Southwest Philosophy Review*, and *Ancient Philosophy* (forthcoming) and is completing articles on Plato's psychology in the *Phaedo*.

**Cass Weller** earned his Ph.D. from the University of Pittsburgh. He teaches at the University of Washington, where he works in Ancient Philosophy, Modern Philosophy, and Epistemology.

**Bernard Williams** is Deutsch Professor of Philosophy at the University of California, Berkeley, and a fellow of All Souls College, Oxford. He has held chairs in London, Cambridge, and Oxford, and from 1979 to 1987 he was Provost of King's College, Cambridge. His publications include *Morality, Problems of the Self, Descartes: The Project of Pure Enquiry, Moral Luck, Ethics and the Limits of Philosophy, Shame and Necessity*, and *Making Sense of Humanity*.